

nature*June 3, 1976*

UN prepares for the next round

It is a safe bet at this time of Habitat that few people, even those concerned with United Nations affairs, have ever heard of the CSTD. The acronym stands for the Committee on Science and Technology for Development. At present, unlike many such international groups, the CSTD provides a forum for debate, rather than the sort of political battlefield which is now all too familiar. But the protagonists are the same: on the one side the countries of the Third World, on the other the industrialised, affluent countries. At the CSTD's third meeting, in New York earlier this year, the debate was concentrated for something like half the 3-week session on a single issue: who should run the United Nations Conference on the Application of Science and Technology to Development, now scheduled to take place, probably in Vienna, in 1979.

Even the proposal for this event has received hardly any publicity, inside as well as outside the UN; but enough information is now available to indicate certain differences from previous gatherings in the long series that started, in effect, with the first Atoms for Peace Conference in 1955. Probably the most important point is that from the start this has been intended as an inter-agency conference, the UN itself playing a central but much smaller role than hitherto. This should have the immediate advantage that the specialised agencies, almost all of which have a stake somewhere in the field of applied science or technology, need not feel that the meeting will be dominated by the political preoccupations of the UN.

At the same time, it is impossible to envisage the preparation for such a conference without continual and probably bitter in-fighting between the specialised agencies themselves—as witness the unheralded personal appearance of Mr M'Bow, the Director General of UNESCO, to ensure that his agency obtained a fair hearing at the recent session in New York. Probably one reason for this was that the organisation principally interested in the conference appears to be not one of those concerned with science as such—WMO, WHO, FAO, ITU, or UNESCO itself—but UNCTAD, whose name is becoming more and more closely linked with the current enthusiasm for the “transfer of technology”; and this, apparently, is to be the real theme for discussion in 1979.

What information is available about plans for the conference indicates that it will follow in a general way the pattern set by the Stockholm environment conference of 1972. This involves a series of pre-conferences at regional level, intended particularly to digest the now-familiar country reports from governments, and a “forum”, more serious in intent than that at Stockholm, at which organisations and presumably individuals not represented at the conference itself can debate the approaches to science and technology most likely to be of use to the Third World.

Two innovations will be welcome to those who have felt that the dead hand of the UN Secretariat predisposes all such meetings to argument rather than action. One is the CSTD's insistence that the conference secretariat should be seconded from the agencies directly involved in the subjects to be discussed. The other is that the conference secretary-general should be selected by the CSTD, being the body most closely concerned, and not presented as the personal choice of the UN's Secretary General, as has hitherto appeared to be the case.

It was this last issue that occupied a whole week of participants' time in New York. At first, it seems, Third World representatives insisted that the secretary-general should be chosen from among their own eminent scientists or technologists. In this they were backed by a number of other countries, notably, Australia, Canada and Sweden. When it became apparent that there was just no suitable candidate available from within this group, it was agreed to accept one from outside.

There is, as it happens, one widely canvassed name: Guy Gresford, who was in fact Australia's representative at the meeting. As former secretary of the UN's Advisory Committee on the Application of Science and Technology, he is known and respected by governments and UN specialised agencies alike; he is acceptable to the Third World majority group; and as a quiet, confident administrator he is as likely as anyone to be able to see fair play in the battle already being joined between the various members of the UN family. They are now for the first time embarking on a major inter-agency conference; its success will depend largely on their ability to work together for the benefit of the countries whose aspirations they so often claim to support. □



Give yourself a write-up

ONE of the reasons that is regularly given for the decline in numbers of students wanting to study natural sciences and the concomitant growth in the numbers of those wanting to study, in particular, medicine and law, is that whilst medicine and the law is about people, science is about things and is therefore less human. Although the argument about the people-orientation of the law is one that is open to considerable debate, it is more important that the argument that science is about things is not allowed to pass unchallenged. For if science cannot present itself as a human activity, it not only loses potential bright students but it also positively encourages a science-for-people movement which at times seems only too glad to present science as practised by heartless technicians.

There have, of course, been a few who were prepared to say that science is only what its practitioners can agree amongst each other and thus does not reside ultimately in textbooks or learned journals. This puts science firmly in the realms of discussion, persuasion, criticism and dissent—very human activities. But scientists have rarely helped along this cause by leaving for others a very helpful record of their intellectual development, and if pressed have tended to regard scientific publications as adequate evidence.

At the same time, the historians of

science seem to have increasingly fought shy of the twentieth century, not just because many of the major figures are still around (this, after all, has not inhibited historians in general) but also because to write on the recent past without putting it into a social, even ideological framework is somehow more difficult, although no less necessary. And social and ideological frameworks have become so much the fashion.

If there are a few very distinguished exceptions, such as Margaret Gowing both with her histories of atomic energy and her Contemporary Science Archives in Oxford, these only point to the need for more. If historians take too much refuge in Newton and sociologists in counting references, scientists should step into the breach themselves.

The web of personal relationships, ideas discarded, intellectual growth, differences of opinion in science is perhaps less fascinating to the general public than the revelations of a film-star or even a vet. But it is certainly not worthless, and it can only really be relayed at first hand—by autobiography. Many published autobiographies are deceptive, self-serving or tedious but the accumulation of them helps a clear picture to emerge over the course of the years.

Most writers of autobiography have more than simply a desire that their

world-view should be widely known; they are used to making a living by writing and view an autobiography as one more way of keeping the wolf from the door. The scientist does not live under the same sort of shadow and perhaps this is one of the reasons that scientific autobiography is rare. Another may well be that the skills, in recent years, of a Waugh, Russell or Muggeridge positively inhibit non-literary characters from putting pen to paper. As a result, however, we are sadly deficient in any personal records of some of the scientific developments of the past fifty years—what people used to do, think, read, be influenced by, care about—and biography is not sufficient to remedy this (nor for that matter is scientific biography all that frequent at the sort of depth which helps one understand the human side of science).

Could not more scientists, including those with no particular eminence to their name, be persuaded to spend their last year or two in a tenured and secure post prior to retirement in getting down on paper—if only for it to sit in a university archives—what doing science was really like for them? This is something that the Royal Society could well stimulate, regardless of questions of literary merit. And who knows, there may even be a few more Watsons, Snows, Medawars out there. Well, one or two. □

The human breeding strategy

Paul A. Colinvaux, Professor of Zoology at Ohio State University, takes issue with what Nature in a recent editorial called "the now widely accepted view" that high population growth is a consequence and not the cause of poverty

THE VIEW that high population growth is a consequence and not the cause of poverty is widely held by demographers, but it is not held by ecologists interested in breeding strategies of Darwinian species. The breeding strategy of the species *Homo sapiens* is comparable to that of many birds: a few large young are produced at a time, then fed and protected for a long

period. This breeding strategy, which I have called "the large young gambit", confers fitness by yielding a high return of surviving and reproducing young for each unit of resource capital invested. The advantage of being given the large young gambit by natural selection can readily be seen by contrasting it with the breeding strategies of species which make very large num-

bers of small eggs or seeds, what I call "the small egg gambit".

Small eggs, and the tiny young which hatch from them, are at high hazard, so that most of them die. Since every tiny corpse represents a unit of resource investment upon which there is no return, simple economic arguments show that fitness is likely to be greater when a small number of large young are reared and protected than when multitudes of tiny progeny are cast out to fend for themselves. The breeding strategy of people, therefore, is one that gives them great ability to

maximise the return, particularly in circumstances of limited resources. It is a property they have in common with all followers of the large young gambit, with sparrows, antelopes, viviparous snakes, and great white sharks.

David Lack showed us 20 years ago that an essential requirement of the large young gambit is that the size of the family be closely gauged to the available resources. He produced good evidence that the clutch size of many birds is dependent on the food supply, the number of eggs laid varying from season to season. The original data can be found in his *Natural Regulation of Animal Numbers*, and there has since been overwhelming confirmation of his work. The mechanism deciding the clutch-size for a species following this gambit must ensure that the chosen clutch is not slightly too small for the available resources, or the genes for this mechanism will be removed from the population by natural selection favouring those genotypes which gave their possessors a slightly more ambitious clutch. But it must also ensure that the clutch is not too large, because an attempt to rear a family too numerous for the available resources puts all the progeny at hazard. The most fit individuals will be those which assess the food and other resources of their living space with the most precision.

Natural selection was undoubtedly alive and well when the breeding strategy of the Pleistocene populations of *Homo* was fashioned. Our ancestors were, therefore, adapted to make the best possible choice for the number of children in a family, "best" meaning the largest number for which there was a high probability of it being possible to rear them to maturity. There can be little doubt that the workings of intelligence gave people a skill in applying this breeding strategy which could not be equalled by birds, sharks, and the rest. It is against these Darwinian realities that there is scope for reappraisal of what we know of reproductive habit and taboo in primeval societies. Infanticide, for instance, appears to be, not a device for impeding population growth, but rather a device for increasing it. Infanticide is one of the mechanisms provided to human families by natural selection not to make the family too large for the available resources. It increases fitness. The pill, the diaphragm, and the lump of plastic perched in the uterus must be expected to work in the same way. Contraceptives ought to increase Darwinian fitness—a thought for those who ponder the failure of birth control schemes.

It appears that no human population has abandoned the Darwinian breeding strategy of our Pleistocene ancestors.

Every couple continues to choose the largest number of children it can afford to rear with a high expectation of success. The change we have made is to abandon many parameters of the old niche, replacing them with surrogates, and we have added to the niche many new parameters or alternatives. These changes in niche are well understood: farming instead of gathering; herding instead of hunting; concentrating in cities, which allows "agribusiness" and the bringing of food to people instead of people to food. These changes in niche are the real measure of the difference between people and the rest of Darwinian species, none of which can change their ways of life to suit their convenience as we can.

If a niche requires many resources, or allows many alternatives, an ecologist calls it "broad"; if a species is of stereotyped habit and highly specialised, its niche is called "narrow". Humanitarians refer to similar things when they talk of "riches" and "poverty". Rich people have large niches, poor people have small ones. It takes more resources (tangible and intangible) to provide niche space for a rich human couple than for a poor one.

But rich and poor alike retain their large-young breeding gambit. Each couple of both assesses the resources of its habitat in ways which have been programmed for them by natural selection from the times of remotest ancestors. We are conscious of these ways when we know that we love children. And our process of resource-assessment must always have taken into account the traditional ways of life. Even primeval people learned how to live from their parents and their parents' peers. Darwinian fitness (given by success at child-rearing) for these people must have been dependent on providing for each child the resources needed to let it grow up to live in the traditional way, for otherwise it would never be equipped to mate and rear children itself. A couple planning a family estimated, as it still does, the resources needed to raise a child to live as well as the parents lived. The number of children the couple chooses is the largest for which traditional resources can be provided. That people still behave so is the explanation for the fact that families in affluent countries tend to be smaller than families in poor countries, the so-called demographic transition. A well-off couple chooses the most children it can raise to well-off standards or even to slightly better standards, which is not likely to be many. The poor only have to provide for each child what is needed to raise it in poverty, which may allow for more children even when the actual resource base is less.

It is our experience with the demo-

graphic transition which has allowed the claim to be made that high population growth is a consequence and not a cause of poverty. The element of truth which lies in the claim is the observed fact that families of the affluent tend to be smaller, for the good Darwinian reasons described above. But this is not to say that "high" population growth does not happen in the affluent (it can and does—see the history of the United States); it is merely to claim that the highest growth rates of the contemporary world are in poor nations. And it speaks not at all to the question of the cause of poverty itself. This too can be understood and predicted from a knowledge of the human habit of changing niche while keeping breeding strategy constant.

While people retain their excellent large-young gambit they can be expected to maximise the return, and this is true regardless of their standards of life. Increase in population with time will have a slightly higher value for the poor, but this difference in value makes no difference to the size of the eventual population limit. Poverty is, of course, a relative value, and is better replaced in the interests of clear thinking by some measure of resources, or by the concept of niche-size itself, which G. E. Hutchinson showed us could in theory be quantified (see *Cold Spring Harbor Symposium on Quantitative Biology*). If riches and poverty are large niches and small, then they are properties of the size of population at all times except during rapid population growth.

Times of rapid growth are the exception because these must be times of excess resources. This is the explanation of affluence and rising numbers going together in the United States during the last century. There were for a time so many spare resources that almost unlimited choices for life were possible, giving rise to the most extravagant portions of the American Dream. But as numbers close on resources, as the population presses on the population limit, choice must go. Then, if the numbers continue to rise, the resource allocation per individual (its niche space) must decrease. People can accept changed niches, it is their unique quality. They continue to maximise the return even as the upper limit for their traditional way of life is reached. The next generation must respond by reducing the size of the niche so that the limit can be increased. But they, in turn, still maximise the return. All poverty is caused, in the long run, by the continued growth of population. The continued growth of population itself is caused by the combination of our excellent Darwinian breeding strategy and our ability to change the niche within wide limits. □

Iranian transplant

Iran is in a hurry to industrialise, and needs teachers and technicians. **Peter Newmark**, just back from Tehran, assesses the problems

EVEN a brief visit is enough to show that the development of science and technology in Iran is hurtling forward, somewhat blindly, along the paths well worn by the Western countries. Using but a tiny portion of his oil revenues, the Shah (or, more properly, the Shahanshah or King of Kings) has instigated an immense programme of higher education to produce the technologists and scientists that his country needs to hoist it from "developing" to "developed" status. But all the oil in Persia cannot make the hoisting operation run smoothly.

In other circumstances a country with an illiteracy rate of 50% would focus its attention on a programme of general education at the initial expense of higher education. Iran, however, is in a hurry, and under few enough financial restraints that it can undertake parallel improvements at all levels of learning. The Shah has therefore recently introduced free compulsory education to the age of 13 and, last year, added free university education in return for a commitment to join the civil service for a period of twice the length of the degree course.

The consequent demand for higher education has greatly exceeded predictions. This year there are eight applications for every university place that is available, in spite of the fact that the number of places has doubled in less than five years. In part the additional demand will be met by the introduction this year of a Free University on the radio and television networks (based on and bought from Britain's Open University).

The clamour for higher education is a consequence of over-eager policies, particularly in a country where the material advantages that accrue to the highly educated minority can all too clearly be seen. The problem is both to meet the demand and to channel it sensibly. In the first place there is a drastic need to reinvest many of the educated into the teaching system, which at present suffers from manpower shortages in many areas.

The other priority must be to solve Iran's shortage of skilled technicians. That problem is the classic one faced by any country that suddenly finds itself able to buy up the highly advanced technology of the developed countries—be it a nuclear power plant or an electron microscope—but without the trained technicians needed to run

and maintain the equipment. For example, Dr Zarghami, Chancellor of the Arya-Mehr University of Technology in Tehran, estimates that the country presently lacks 2,000 of the 28,000 graduate engineers that it requires. But worse is to come. In ten years time he predicts that Iran will only be able to provide 60,000 of the 75,000 graduate engineers needed.

That problem is essentially insoluble in the short term. It is not even clear that the output of graduate engineers can really be more than doubled in ten years despite the founding of a second campus of the University of Arya-Mehr. The campus, a self-contained city, will be situated a few miles outside the city of Isfahan, 200 miles south of Tehran. When fully functional the Isfahan campus will be the largest technical university in the Middle East with the possible exception of the Technion in Haifa. In ten years time it will have 10,000 undergraduates and 3,000 graduate students.

Another part of the plan to increase the number of technicians is the intention of the Ministry of Science and Higher Education to collaborate with Iranian industry in setting up an Industrial Training Complex. Its sole function would be to turn out specialist technicians to satisfy the requirements of industry. At a lower level there is already a requirement that any factory which employs more than 500 people has attached to it a school for the children of the factory workers. Such schools can instruct the children in at least some of the basic skills required for them to follow their parents into the factory.

In spite of all these plans the shortage of specialist technicians will still be increasing rather than decreasing for at least the next ten years. The gap will have to be met by offering high salaries to skilled personnel either from the West or from other Middle East countries. Recruitment from the latter, to judge from accounts of the recent Islamic Solidarity Conference in Science and Technology, is bound to provoke regional discontent.

With such a shortage of teachers and technicians it would not be surprising were the progress of basic scientific research to be in jeopardy. On the contrary, however, a clear and strong commitment to the establishment of basic research is evident both in universities and in separate research insti-

Sorry, for copyright reasons some images on this page may not be available online

Cam & Pen Int.

Iranian village

tutes in Iran, although fears are expressed in some quarters that the level of future support will depend very directly on the sales of oil which, together with the revenue thereby derived, fell sharply last year.

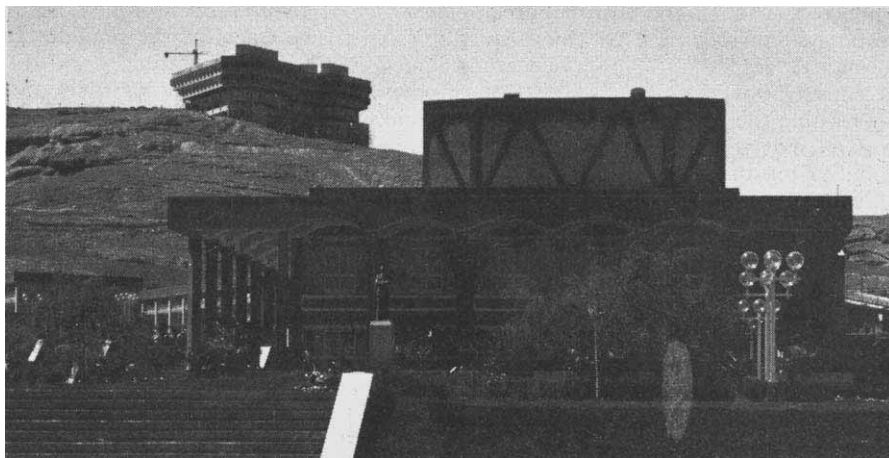
In the meantime the administration of research is very much along Western lines. For example, last year saw the creation of the National Scientific Research Council, a central policy body. The council is responsible for the co-ordination, supervision and funding of research. The latter will be subject to the usual kind of peer group approval and will take into account the council's suggestion of priority areas for research. At the moment, however, even Dr Parsa, Deputy Minister for Science and Research, admits that priorities are largely hypothetical since any scientist who is of reasonably high calibre (and who is at least not openly critical of the Shah's policies) is likely on those grounds alone to be funded. That situation is likely to remain for some considerable time until the number of good research scientists reaches a critical mass. Demand will then no doubt start to outstrip the supply of funds.

Until (and to ensure that) the critical mass is reached, it is essential that Iranian scientists undergo part of their training overseas and that international contacts are lovingly fostered. At present about 50,000 Iranians are overseas partaking of some form of higher education. Many science graduates like to obtain their PhD overseas, but Dr Mofidi, the Vice-Chancellor of Tehran University, would prefer to see overseas training at a post-doctoral level so that it can be based on a genuine familiarity with what awaits the scientist on his or her return. And return he or she almost certainly will, given the very favourable job market

in Iran compared to the West. That may also enable Iran to buy up a number of Western scientists to tide them over the next few years.

International contacts are necessary not only for the placing of students overseas but also for the development of research programmes. Contacts are encouraged through sabbatical leave every fourth year and by financing international meetings in Iran. A recent example was the four-day meeting on the eukaryotic genome at the University of Tehran in May. About 40 Western experts were imported to expound on their research before an audience of about the same number of local scientists. No doubt the prestige of such an occasion is as important as the chance of making contacts; the actual symposium presentation probably come a poor third in real value.

Whereas Iran is presently looking determinedly to the West for international collaboration and advice, many Iranian scientists would like to develop a more regional approach. In particular that would make sense for programmes of "regional technology" in, for example, solar energy, desalination and arid zone research. Already there are programmes of cooperation with various Arab countries, Turkey and Pakistan. A bigger attraction for some would be Israel because of her scientific expertise both in basic research and in "regional technology". Although there is already a certain



University of Shiraz, Iran

amount of informal contact with Israel, political factors rule out any real cooperation.

There is in Iran an ancient tradition of science and technology. Its origins lie in the precise requirements of the Islamic faith for the time and direction of prayer. But tradition has nothing to offer to the new generation of Iranian scientists who are whole-heartedly committed to transplanting Western science, warts and all, to the Middle East. Whilst recognizing the need for a scientific discipline to have roots in the society that supports it, they argue that Iranian society is in the process of general change and in time will be ready to nourish the roots of Western-style science. The concurrent argument

is that without generous support of basic research, no roots, in the form of Iranian eminence, will form.

The transplantation of Western values in science and much besides has, not surprisingly, led to certain problems in the universities as well as in Iranian society at large. In particular there are difficulties of adjustment for the growing number of students who come from poorer, more traditional backgrounds. The result is a high drop-out rate and sporadic political dissent which is ruthlessly suppressed. Overall, the Shah will not for the present tolerate any public questioning of his drive to make Iran a military and industrial force to be reckoned with. □

Conservation in Iran

Peter D. Moore, recently in Iran, looks at attempts to conserve wildlife in a fast changing country

NATURE conservation in the industrialised nations is said to be a cause associated primarily with the middle classes. Developing countries often see it as a worthwhile economic exercise, since wildlife parks can be a major tourist attraction providing foreign currency. But in Iran it is neither of these. There is little tourist industry apart from that which centres upon the archaeological rather than wildlife potential of the country. Yet conservation ranks high among the priorities of the government. The Iranian Department of the Environment, headed by Eskandar Firouz, is (unlike the British establishment of the same name) primarily concerned with environmental research and practical management problems in Iran's cities, rural areas, wildlife parks and reserves, and around its Persian Gulf and Caspian Sea coasts.

The sheer size of the country (about three times the area of France) and its

altitudinal variation (from 50 m below sea level at the Caspian to 5,700 m in the Elburz mountains near Tehran) results in a wide range of habitats and a high diversity of wildlife. About 80% of Iran is over 1000 m in altitude: much of this lies in a great central plateau surrounded by high mountains, the physical isolation of which could account for the high proportion of endemics in the Iranian flora.

To ensure that examples of all the major habitats are conserved, the Environment Department has established over 40 wildlife parks and protected areas, covering a combined area larger than Switzerland. The broad-leaved forests of the Caspian area still contain wild pig, brown bear and the Persian subspecies of fallow deer, which has been rescued from the verge of extinction. Tigers were reported from this region in the first half of the century, but are now feared extinct.

The Caspian lowlands also contain some important wetland areas which serve as winter feeding grounds for large populations of wildfowl that breed in western Siberia. Other wetland areas in Iran are mainly in the west, especially the saline lakes: the largest is Lake Rezaieyeh, which extends over 3,000 ha and supports breeding populations of 25,000 pairs of greater flamingo and 1,400 pairs of white pelican. The wetlands of the Gulf in the south of Iran include the Eastern Mesopotamian marshes and the rivers of Iranian Baluchistan, where the marsh crocodile still survives at the western limit of its range.

The wealth of wetland habitats which Iran contains, coupled with the threat which is posed to them by increasing industrialisation, may explain why Iran has become a global trend-setter in the field of wetland conservation. In February 1971, a Convention was agreed at an international conference in Ramsar, Iran, by which all signatory member countries should designate at least one of their wetland sites for inclusion in a list of wetlands of international importance. Iran was the second country to ratify this treaty,

while it has taken the British bureaucratic system five years to achieve this formality, despite the lip-service paid to conservation.

Probably the greatest threat to Iranian wetlands is posed by the industrial pollution of the Caspian Sea and the oil pollution of the Persian Gulf. In the Caspian much of the pollution is derived from the Soviet Union, but there is now cooperation between the two countries in trying to reduce its level. Worst affected is the fishing industry, where the production of sturgeon and caviar provides an important export industry. Attempts to reduce pollution are being coupled with the building of fish breeding stations. Near Rasht on the Sefid river, which runs into the Caspian, one such site of 132 ha has been built with Soviet aid and at a cost of \$5 million. It produces 4 million fish per year which are released into the river.

At the other end of the habitat spectrum, almost half of the surface area of Iran is covered by desert or arid steppe, and obviously this type of habitat is well represented among its reserves. Here, in particular, human exploitation of the environment and wildlife conservation reaches a delicate and precarious balance. A relatively small percentage of Iran's population is involved in pastoralism and these people are found mainly in the semi-arid areas. Precise and accurate figures are difficult to obtain, but it has been estimated that about 400,000 pastoralists maintain a total of 30 million sheep and 15 million goats, so the grazing pressure over the arid lands is extremely high. These herds reduce the wildlife potential of the areas they graze, but it is difficult to establish whether this is a result of their consumption of herbage or simply the disturbance caused by a human presence.

The arid areas contain a rich and varied assortment of wildlife, now much depleted through hunting. Wild ass, goitered and jebeer gazelle, cheetah, leopard, wild sheep, great and houbara bustards, all still survive in viable populations within the protected regions. The large raptorial birds, particularly the golden eagle, are surprisingly frequent given that such species have often suffered at the hands of pastoralists. The maintenance of this wildlife must depend, however, upon restricting any further human pressures upon their dwindling populations, and this raises a number of social and political problems.

Before any such problems can be solved, it must be ascertained precisely what form man's interaction with wildlife takes. Studies of the productivity of the natural flora are needed urgently, particularly in the *Zygo-*

phyllum/Artemisia scrub which occupies extensive tracts of the semi-desert areas. The feeding preferences and the potential consumption of the wild and domestic herbivores must also be known, so that competitive interactions can be identified and the carrying capacity of the environment for each species determined. Also, in spite of Iran's wealth of fossil fuel deposits, there is an acute shortage of fuel in the settlements of the desert area and much shrub material is gathered for combustion. A knowledge of the social practices of the people is therefore needed, including an understanding of the movements of the small groups of pastoralists who indulge in transhumance, leaving the desert plateau for the mountain pastures each spring. The historical and archaeological perspective is also important, for some desert areas have been denuded of scrub, and perhaps even of pistachio woodland, for charcoal burning and smelting in the past. It is possible that the great salt lakes or *kavirs* will contain preserved, stratified pollen from which past changes in vegetation can be determined.

An integrated study of these various aspects of the desert ecosystem is currently being undertaken by the Department of the Environment's Bureau of Arid Lands Ecology in the Turan Protected Area of eastern Iran. The solution to another environmental problem, dune stabilisation, is also the subject of an experimental approach. Some 5 million hectares of sand dunes, mainly in south-eastern Iran, provide a very unstable substrate which results in low plant cover and productivity. Since 1968, Esso Research Ltd has been spraying dunes with oil to provide a temporary stabilization of surface sand whilst plant life is becoming established. Oil is deposited at a density of about 4 tons per hectare, and the area is then planted with fast-growing shrubs, some of which may attain a height of 2 metres within 2

years and seem unaffected by the oil deposits around their roots. Although the initial effect is unsightly, the final stabilisation is said to be more effective than that achieved by the old system of using brushwood palisades and, in a country where oil is cheap, the method may provide a fast and efficient method of dune reclamation, whether for human or wildlife exploitation.

Plans are meanwhile being put into operation to develop an environmental park, called "Pardisan", within the city of Tehran. A 300 ha site will eventually contain museums, planetarium, aquarium, laboratory complex and a series of reconstructions of various habitats, including fauna, from Iran and other parts of the world. As with the Environment Department's other conservation activities, emphasis is to be placed upon the role of man in the ecosystems portrayed, so the zoological/botanical garden is to be combined with a kind of folk museum in which whole villages will be constructed typical of such areas as the Persian Gulf and Baluchistan. This grand conception, the brainchild of Firouz, has much about it which is attractive, but its practical feasibility remains to be demonstrated.

The one pervasive feature of the Iranian approach to wildlife and habitat conservation is its insistence upon the inclusion of man as a component of all its ecosystems. This country, with about 10,000 years of agricultural history, bears everywhere the mark of man's activities. There are regions which almost certainly could be made more productive if man were to reduce his demand on the environment, but no realistic conservation policy can ignore or totally exclude his presence. It is to be hoped that such enlightened attitudes will penetrate to all levels of Iranian society and will result in the achievement of a balance from which all organisms, including man, will ultimately benefit. □



Date palms on northern edge of Great Kavir

USA

California's nuclear war

Californians will shortly give their view on nuclear power. Colin Norman reports on a critical referendum

WHEN California voters work their way through their cluttered ballot papers in the June 8 primary election, they will decide more than the political fate of a handful of Presidential aspirants. They will also determine the future for nuclear power in California and, perhaps, in a number of other states as well.

A proposition, placed on the ballot by petition last year, would set such stringent controls on the construction and operation of nuclear power plants in California that, if passed, it could bring the nuclear industry to a grinding halt there. The outcome of the vote, the first ever plebiscite on nuclear power, is considered so important to the nuclear industry that, together with various supporting groups, it is pouring millions of dollars into a massive campaign to defeat the proposition. California's nuclear war, in fact, is even overshadowing the crucial battle between would-be Presidents for the state's delegates.

The vote will have repercussions extending well beyond California's borders, which is one reason why the matter has assumed such importance. For one thing, voters in several other states will be faced with the same choice in the November Presidential election. Similar propositions have already qualified for the ballots in Oregon and Colorado, they are also considered likely to qualify in North Dakota, Wyoming and Washington, and anti-nuclear groups are hoping to force nuclear referenda to be held in a number of other states next year. The theory is that a victory in California would give a huge boost to the anti-nuclear movement in those states. But, equally important, if the California proposition passes, it could undermine political support for nuclear power in Washington and in state legislatures around the country.

In fact, the proposition's supporters can already claim a victory of sorts. Three bills, considered milder versions of the proposition, are sweeping through the California legislature and are expected to be approved before the primary election. The bills were undoubtedly snatched by the huge publicity given to the campaign, and they would probably never have stood a chance if the proposition had not gathered so much momentum.

Even at this late stage, the outcome of next Tuesday's vote is in some doubt. Although public opinion polls have consistently given opponents of the proposition the edge, the latest poll, conducted in mid-May, indicated that 29% of the voters were still undecided. Of the others, the poll found 30% in favour, and 51% opposed. Those figures are one reason why both sides are conducting a fierce campaign in the final days before the June 8 show-down.

The proposition, and the campaign which is now swirling noisily around it, is an interesting exercise in participatory democracy, and an object lesson in the wild ways of California politics. It will be Number 15 on the ballot papers, buried among questions on issues ranging from bank taxes to legalised bingo.

Formally known as the California Nuclear Safeguards Initiative, Proposition 15 would set the following conditions on the construction and operation of nuclear power plants in California:

- In the event of a nuclear accident, there must be no limit to the total damages which could be claimed by victims. At present, a federal law, the Price-Anderson Act, limits the liability of the nuclear industry to a total of \$560 million per accident.

- Within five years of passage of the proposition, the state legislature must have determined, by a two-thirds vote, that nuclear reactor safety systems will function correctly. Such a determination would be made on the basis of tests of the systems conducted "in actual operation".

- Similarly, the state legislature must have determined, by a two-thirds vote, that radioactive wastes can be stored without adversely affecting the health of Californians of their environment.

Unless those conditions are met, no new plants could be licenced and existing plants would be able to operate initially at only 60% of their rated capacity. Their output would then have to be decreased by 10% a year until they are phased out in the mid-1980s.

The proposition qualified for the ballot early last year when its supporters gathered together the necessary 312,000 signatures. The early sponsors of the proposition were a San Francisco-based group called Californians for Nuclear Safeguards and a Los Angeles organisation known as Peoples' Lobby, but much of the original impetus came from Ralph Nader's organisation. An important addition to the pro-initiative forces was

the Creative Initiative Foundation, a religious-philosophical group whose lobbying arm, known as Project Survival, joined the quest for signatures to qualify the proposition early in 1975. Three nuclear engineers who quit their jobs at the General Electric Corporation last January to campaign for the proposition were all members of the Creative Initiative Foundation.

Opposition to the proposition is spearheaded chiefly by a committee of representatives from the nuclear industry, businessmen and labour leaders, called the California Council for Environmental and Economic Balance. Headed by Edmund G. Brown Sr, former Governor of California and father of the present incumbent, the committee is well-financed and has hired a firm of campaign consultants to manage its effort.

Because there has been some confusion among Californians stemming from the fact that a 'no' vote on the proposition means a 'yes' for nuclear power, the committees recently changed their names. Supporters of the proposition now campaign under the banner of the "Yes on 15 Committee", while their opponents are headed by the "No on 15 Committee".

The rules, and perhaps the outcome, of the campaign were abruptly changed last January in favour of the opponents of Proposition 15. In the 1974 elections, California voters approved a proposition, again placed on the ballot by the Peoples' Lobby, which would set strict limits on the amounts of money spent by candidates in state-wide elections. But in January, the California Supreme Court ruled that the limits do not apply to campaigns for referenda, and money then began pouring into the coffers of the opposition groups, much of it coming from out-of-state utility and energy companies. By the time the vote is taken the various No on 15 groups will have spent about \$7 million, according to some estimates. By contrast, supporters of the proposition will have amassed only about \$800,000. Joyce Koupal, who with her late husband Ed Koupal formed the basis for People's Lobby, claimed last week that the opponents are "buying the election".

In the last week or so of the campaign, the No on 15 Committee will sponsor a barrage of television and radio commercials, pushing the argument that the elimination of nuclear power in California will cause unemployment to rise and living standards to fall. And the Yes on 15 Committee will have Ralph Nader stumping the state, arguing that energy can be conserved and alternative sources tapped, in such a manner that living standards

will not suffer. Nobel laureates have already been ranged on both sides of the issue, and arguments have been traded back and forth about nuclear safety, safeguards and so on. The California campaign, in short, closely reflects the debate about the merits and hazards of nuclear power which has been raging nationally for the past few years.

But there is one additional issue in the campaign—the problem of operating nuclear reactors in a state where the seaboard is criss-crossed by geological faults, and the interior is almost devoid of cooling water. Because of those factors, experience with nuclear power in California has so far been vexing for the nuclear industry.

There are now three plants in operation, four under construction, and some 20 more at various stages of planning. Of the operating reactors, one small unit located north of San Francisco is now being reviewed by the Nuclear Regulatory Commission because an active fault has been discovered close by, while another has been plagued with nagging technical problems ever since

it began operating early last year. A storm of controversy has recently arisen over two reactors under construction in Diablo Canyon, midway between San Francisco and Los Angeles, because an offshore fault has been discovered some 3 miles from the site. And it can safely be assumed that similar site problems will arise for future reactors.

Supporters of Proposition 15 thus have a plump target to shoot at, but the opponents are also not short of big guns. One of their chief arguments is that California already imports some 22% of its energy from the other states, and that without nuclear power its dependence on other regions will increase greatly. A substantial source of imported energy at present is a series of dams on the Colorado River, and a number of coal-fired power stations in Utah. But the Colorado is already dammed so heavily that it no longer reaches the sea (it disappears into the sand in Mexico), and Utah residents are growing reluctant to put up with air pollution in order to supply Los Angeles with power. A few weeks ago,

for example, the Southern California Edison Company cancelled plans to build a massive coal-fired plant in Utah because of environmentalist opposition. Opponents thus argue that without painful conservation measures and greater reliance on imported oil, California would not be able to meet its energy needs, and it will suffer economically as a result.

One factor clouding the picture in the late stages of the campaign concerns the effect that passage of the three nuclear energy bills by the state legislature will have on the prospects for Proposition 15. The bills, which deal with nuclear safety, safeguards and waste storage (but not with liability limits) would require the recently created California Energy Resources, Conservation and Development Commission to determine whether the federal government has dealt satisfactorily with nuclear power problems, and the commission's findings would have to be approved by a simple majority of the legislature. Unless such approval is given, no new reactors could be built, but the legislation would leave untouched the three existing plants and the four under construction.

Another uncertain factor is whether or not Proposition 15 would survive a court challenge. If the proposition is passed, the nuclear industry would be certain to challenge its constitutionality, arguing that it would usurp the federal government's power to set regulations for nuclear power. Legal experts are divided on the likely outcome of such a challenge, but two Stanford University law professors who have studied the matter conclude that most of Proposition 15's provisions would just survive a court test.

If, as now seems likely, Proposition 15 will be defeated by a small margin, it is likely that both sides would claim a victory. In a speech delivered on May 4 at the Forum Atomique Européen in Madrid, Carl Walske, president of the Atomic Industrial Forum, predicted victory for the opponents of Proposition 15, and claimed that such a result would sound the death knell of the anti-nuclear movement. The campaign, he said, "may well be their last big drive".

But if supporters of Proposition 15 come even close to victory, they will claim success, arguing that they were defeated by the industry's money and that the extent of the nuclear opposition is such that there is no mandate for a vast expansion of nuclear power in California. As a spokesman for the Yes on 15 Committee put it last week, the industry knows that it must defeat the proposition heavily, because anything short of a landslide defeat would be called a victory by the anti-nuclear groups. □

Earthquake money

A large helping of money for seismology, together with the establishment of a top-priority programme to reduce the destructiveness of future earthquakes, is promised in a bill approved without debate last week by the US Senate. Proposed by Senator Alan Cranston, the senior Senator from California, the bill has been hovering in the congressional wings for nearly four years; it was rushed to centre stage and swiftly passed last month in the wake of the destructive earthquakes in Italy and the Soviet Union.

If it survives the rest of the congressional process intact, the bill would authorise expenditures of some \$150 million over the next three years on a massive programme of research on earthquake prediction, efforts to improve the resistance of buildings to major shocks, studies of the social and psychological consequences of earthquakes, and how to issue earthquake warnings without adding to the disaster.

The bill would also establish a top-level committee of 15 people, appointed by the President, to keep tabs on the effort and to advise the President on its progress, and it would require the White House to prepare annual reports evaluating progress in reducing earthquake risks. Those two provisions would give the effort considerable political visibility. The funds would be shared mostly by the US

Geological Survey, which would be authorised to spend \$75 million in the next three years, and by the National Science Foundation, whose share would amount to some \$60 million.

Describing the measure last week, Senator Cranston noted that 39 states are susceptible to major earthquakes. He also pointed out that a major earthquake in a metropolitan area in California would claim thousands of lives and cost billions of dollars in property damage. "We may be courting certain disaster if we fail to take preventive measures now", he said.

The money is, however, far from being safely in the bank for seismic research. To become law, the bill must also be approved by the House of Representatives before the end of this year. Key legislators there have said, however, that they will hold hearings on the measure as swiftly as possible. The Ford administration is on record as opposing the measure, chiefly because it insists that earthquake research is already proceeding with sufficient priority. Most important, even if the bill is passed and signed into law, the money would still have to go through the annual appropriations process before it can be made available. Nevertheless, the electoral appeal of the measure is obvious and the time is probably ripe for the bill to be approved.

USSR

The clamp-down continues

Instances of harassment of Soviet scientists are still coming to light. Vera Rich reports on recent developments

SOVIET "refusnik" scientists dismissed from their posts after applying for a visa for Israel have recently been threatened with arrest. An appeal to the world scientific community, signed by Professor Mark Azbel, Viktor Brailovskii and others, states that some 20 scientists called in to regional police headquarters in Moscow were told that if they did not commence working within two weeks they would be charged under the "parasitism" law, which covers such offences as prostitution, drug addiction, and alcoholism.

The illogicality of this demand—it is the authorities who prevent the refusniks finding work—was compounded by the fact that it was also made of the few would-be emigrants, notably Academician Veniamin Levich, who have not yet been dismissed. This suggests that this particular threat is directed at all those scientists who take part in the informal seminars which the refusniks, including Levich, hold to maintain some kind of scientific life and exchange of ideas. For some time, uniformed and plain-clothes police have been making occasional visits to the seminars; these visits have recently been stepped up, and on one occasion a lecturer was interrupted with a demand for identity cards. One seminar, in Kiev, has been closed down by the authorities.

Some idea of the authorities' thinking is also to be found in the recent decision concerning cyberneticists Viktor and Irina Brailovskii. Their original visa applications in 1972 were denied on the grounds that Viktor had had access to classified information. Recently, however, his former superior,

Professor Aleksandr Lunts, was allowed to emigrate to Israel; from there he told the Soviet authorities that since his own security status did not preclude emigration, the same should be true *a fortiori* of his former subordinate. Viktor Brailovskii has now received clearance, but Irina has been refused, and has been informed that this is specifically on security grounds. None of her work at the Moscow University Cybernetics Institute was classified, however, as its publication in open journals proves; and since her dismissal in September 1972, there is no obvious way that she could have acquired secret information.

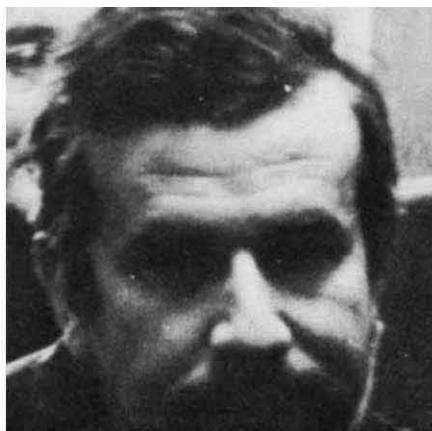
One possible explanation of this anomaly is contained in a recent statement by Al'bert Ivanov on behalf of the Central Committee of the Communist Party of the Soviet Union. Following complaints that the human rights clauses of the Helsinki agreement were not being implemented, Ivanov gave an interview to six prominent activists, including the scientists Mark Azbel, Vladimir Slepak, and Anatolii Sharanskii. Ivanov said that the "terms of 'secrecy' will themselves remain secret. Things which were not secret yesterday might become secret today, and vice versa". Challenged by Azbel that this implied that someone never connected with secret information might be refused a visa in case his work might possibly become secret in the future, Ivanov replied that this was "possible in principle", since each case was decided individually "in the interests of the State"; he further said that there were not, nor would there ever be, any rules on what constitutes "secrecy" or on how long it takes for such "secrecy" to expire.

In their appeal, the scientists stress that pressure from world opinion can effect at least a temporary relaxation of the authorities' attitude: "Our fate

and the fates of our families are in your hands, in the hands of our overseas colleagues", they say. Protests abroad, however, can produce retaliatory action, as one recent case has shown.

Professor Brian Spalding, Chairman of the Editorial Board of the *International Journal of Heat and Mass Transfer*, had been a welcome guest at the first four All-Union Conferences on Heat and Mass Transfer held in Minsk. Since the last Conference in 1972, however, he has been active in the campaign for Academician Veniamin Levich. In November of last year, Professor Spalding received an invitation to the Fifth Conference held last month, and submitted a paper which was accepted. Academician Levich, however, who developed the now standard electrochemical method of determining heat and mass transfer coefficients, was not invited, and on hearing this Professor Spalding protested to the organising committee of the conference. Academician Levich did, finally, receive an invitation, too late to submit a paper; but Professor Spalding was not given a visa to attend, nor was his paper included in the preprints of the conference.

The Fifth Conference was technically an internal Soviet gathering, to which certain foreign guests were invited, and does not technically come under the declaration of the International Council of Scientific Unions on the free circulation of scientists to international conferences. Nevertheless, since Professor Spalding's paper was accepted, the refusal of his visa is being construed as state intervention in the running of the conference. If the case of Professor Spalding was intended as a warning to the world scientific community, it indicates, as Spalding himself has acknowledged, the constant need for a concerted effort to defend the principles of academic freedom, including freedom of movement and communication for all. □



From left: Veniamin Levich, Viktor Brailovskii, Mark Azbel

correspondence

Selling the Sun

SIR,—It is always welcome when, having written something, others take the time to read and criticise it—positively or otherwise. In your editorial (May 20, page 177) on the solar energy report¹ (to which I was a contributor) you wrote an incisive critique which pointed out some weak points but which I think did not sufficiently stress some of what may be its positive aspects.

We did formulate recommendations on priorities in solar energy research and development—see Chapter 12 of the report, from which the table below is taken.

In drawing up these priorities we studied how the USA, Japan, India, Germany, France, and the EEC had done similar exercises, and then took into account potential use within the UK and the export opportunities—the latter we consider very important in view of the already existing overseas markets and the UK's dependence on export income. Thus it is somewhat unfair to say that we advocated "an advance-on-all-fronts" attitude based only on scientific statements. As to finding figures to back up our claims, we certainly did have problems in obtaining the most up-to-date energy statistics and economic predictions. We did our best in the absence of professional economists and government experts—there are, however, 382 references in the report, and we tried to include the basis for all our facts and assumptions (in marked contrast to many other documents which form the basis for important policy decisions).

Futurology is fraught with dangers. Two government examples in the energy field may suffice. In 1952 it was

stated² that since selenium solar cells were only 0.53% efficient, "we do not consider that it is necessary to give any special encouragement to the improvement of photovoltaic cells for power production from solar energy". Silicon solar cells are now 16% efficient and we have to import all our requirements for finished polycrystalline and slices for fabrication. In 1965 the CEBG ordered an Advanced Gas-cooled Reactor (Dungeness B) on the assumption that it would be delivering electricity in 1971 costing 0.457 old pence per unit (note the three decimal point accuracy)³. In 1973 the cost of electricity from an AGR had risen to 0.57 new pence per unit. Notwithstanding these calculations Dungeness B has yet to deliver any electricity (now costing about 2.0 new pence per unit to the consumer) since it won't be ready until 1978!

Solar energy is not "free" as some may claim. It has problems of collection, storage and transmission—as do most other energy systems to various degrees. The present economic accounting systems discourage investments with high initial costs that provide long term benefits—the case for many important solar energy systems. Detailed systems analyses need to be done to help decide the options for solar energy development—the International Institute for Advanced Systems Analysis in Austria⁴ is an excellent example of such work which is immediately applicable to the UK. The construction of demonstration houses, offices, schools and solar equipment, combined with positive incentives to industry and individuals, is the way of the USA⁵. Japan has "Project Sunshine" with defined budgets and objectives out to the year 2000. The UK

needs a combination of these approaches if we are to have alternative viable energy options in the next century to be "sold" here and abroad—and not to buy from abroad.

D. O. HALL

Department of Plant Sciences,
King's College, London

¹ UK Section Intl. Solar Energy Society, *Solar Energy: a UK assessment*, (21 Albemarle St., London, 1976).

² Bullard, E. C., et al., *Research*, 5, 522–529; 1952.

³ Patterson, W. C., *Nuclear Power*, 223–224 (Penguin, London, 1976).

⁴ Häfele, W., et al., *Second Status Report of the IIAA Project on Energy Systems 1975*, Intl Inst. Advanced Systems Analysis, 2361 Laxenburg, Austria, Report RR-76-1; and *Energy Strategies*, Report RR-76-8; 1976.

⁵ Morse, F. H., and Simmons, M. K., *Ann. Rev. Energy*, 1, 131–158; 1976.

Human anatomy

SIR,—Anthony Harris (Correspondence, May 6, page 10) observed that right limbs tend to exceed the left in various statistics and suggests that this asymmetry is likely to have its origin in the greater use of muscles associated with right limbs. There is evidence, however, that the superiority of the right side is already established in the foetus.

The total ridge count of the finger tips is a quantitative trait which is laid down in foetuses during the third and fourth months of gestational age and thereafter remains unaltered. Although the major component of the ridge count is genetically determined and environmental influences are thought to play only a very minor role, there is a consistent bimanual difference, the mean ridge count on the right hand being higher than that of the left (Holt, S., *The Genetics of Dermal Ridges* (C. C. Thomas, Springfield, 1968)).

An early asymmetry is also seen in the size of the foetal testes and ovaries. The right gonad is larger and contains more cells than the left in foetuses aged eighteen weeks. Although the cause is not known, an explanation in terms of greater use can surely be excluded.

These examples suggest that an asymmetry of development favouring the right side precedes any preference in the use of right sided limbs. It would seem that right-handedness has its origin in developmental anatomy, and that the original asymmetry is subsequently enhanced by differential usage. Cultural factors may increase this difference still further.

URSULA MITTWOCH

Galton Laboratory,
University College London

Suggested proportional allocation of research resources in UK for period 1976–1981 with estimates of 1981 possible market split

Application	Suggested proportion of total funds (%)	Estimated market split 1981	
		% in UK	% overseas
Space and water heating	45	80	20
Solar refrigeration and air conditioning	5	5	95
Other thermal applications especially solar pumps, solar engines, solar thermal electric power, etc.	15	0	100
Photovoltaic applications	20	20	80
Bioconversion to fuels and materials* and fermentation of biological wastes	10	40	60
Photochemistry	5	Too early to predict	

*Traditional food production studies excluded

news and views

Not so wealthy stars?

from M. G. Edmunds

PERHAPS the most controversial issue of recent years in the determination of the chemical composition of stars has been the interpretation of the spectra of the so-called "Super-Metal-Rich" (SMR) stars. An interesting analysis of three K giant stars by Ruth Peterson (*Astrophys. J. Suppl.*, **30**, 61; 1976) may help to resolve the disagreement.

The problem may be outlined as follows. The strength of an atomic absorption line in the spectrum of a star is a function of the physical conditions and atomic abundances in the star's atmosphere. Analysis consists of assuming a model of the atmosphere and adjusting the atomic abundances until theoretical line strengths match the observed strengths. It has usually been assumed in abundance analysis that the structure of the atmosphere of a K giant star is uniquely determined by specifying a temperature (on the basis of the distribution of continuum radiation from the star), an acceleration due to gravity and an overall metal abundance—metals meaning all elements heavier than helium. The atmospheric elemental abundances in a K star reflect the composition of the interstellar medium out of which it formed. Most stars have a metal abundance similar to that of the Sun, yet there exist certain stars with absorption lines unusually strong for their assumed atmospheric physical conditions, implying that they have a considerably greater abundance of metals ("SMR") than the majority of stars. Comparison of the luminosities and temperatures of SMR stars with theoretical calculations of stellar evolution has suggested that some of them are considerably older than many stars of lower, solar metal abundance. Thus the SMR stars would not be consistent with the fairly well established near constancy, or slow increase, of the metal abundance of the interstellar medium out of which the stars are born, after a rapid increase in metal abundance during the formation of the Galaxy. So dispute has arisen over whether the anomalous strength of absorption lines in SM stars really does represent an over-abundance of metals.

What Peterson has done is to pro-

vide evidence that the line strengths of the SMR stars are consistent with a normal abundance of most metals, but an unusual atmospheric structure. She selected three stars of very similar temperature and surface gravity to minimise the difference of atmospheric physical conditions between the stars. One of them is the SMR giant μ Leo and another the normal giant κ Oph. By obtaining high quality image-tube observations of the spectra she was able to compare the strengths of fairly weak (and consequently difficult to observe) lines between the stars. The use of weak lines is advantageous since it very much reduces the influence on the line strength of uncertain non-thermal velocity fields and atomic collisions. In a K star atmosphere the conditions are such that atoms are either neutral or once ionised. Peterson shows that lines which are from the predominant species in the ionisation equilibrium (a function of ionisation potential of the element) imply that the metal abundance differences between the stars are at most a few percent. Lines arising from the minor species in the ionisation equilibrium are very sensitive to temperature, and it is these lines which differ in strength between the SMR μ Leo and the normal κ Oph. Peterson shows that the difference can be simply accounted for by a lower temperature in the outer layers of μ Leo's atmosphere. She also shows that normal abundances and the alteration of atmospheric structure is consistent with the anomalously strong absorption observed in photometric spectral bands which first suggested the existence of SMR stars.

But why do SMR stars have atmospheric structures different from normal stars? The one element that really does seem to be over-abundant in μ Leo is nitrogen, by a factor of about two.

Peterson speculates that the high nitrogen abundance results in an increased CN molecular abundance in the atmosphere. This molecule could contribute a large number of absorption lines in the near infrared, and by the requirement of energy conservation in the radiation field this will result in a lowering of the temperature in the

atmospheric region where the absorption occurs. It remains to be seen if theoretical stellar atmosphere models computed with variable molecular absorption can give quantitatively the required drop in temperature.

There also remains the problem of why the nitrogen abundance is high in the atmospheres of SMR stars. But the excess of this one particular element may prove easier to explain than the previous problem of over-abundance of metals thought to be formed in the extreme conditions of supernova explosion. Nitrogen may be formed in nuclear reactions occurring in the core of moderately massive giants, and it could subsequently be mixed up to the surface. Yet this would imply the reaching of a particular evolutionary stage before a star became SMR. Much work remains to be done to see if all suspect SMR stars are consistent with normal metal abundances except for a nitrogen over-abundance, and to check that such over-abundance can be accounted for by the normal evolution of the stars.



A hundred years ago

WE will leave to other journals the task of criticising the present Exhibition of works of Art at the Royal Academy, and without entering deeply into the question of grouping composition, solidity of painting, *chiaroscuro*, perspective, *morbidezza* of flesh treatment, or aerial effect, we will confine ourselves to a few remarks in a less ambitious key, on those pictures which portray animal life. Of this class there are several important examples devoted entirely to the representation of wild or domesticated animals, with others in which the lower forms of creation play but a slightly inferior part; and in these days when the public taste claims a far more conscientious treatment of the subject than in former times, we may be allowed, without being taxed with unfair criticism to examine how far the respective artists have succeeded in fidelity of execution. from *Nature*, **14**, June 1, 105; 1876.

Membranes at EMBO

from M. S. Bretscher

The 2nd EMBO Symposium was held at Hirschhorn-Heidelberg on May 4-8, 1976, and entitled *The Structure and Function of Biological Membranes*.

THAT the second EMBO Symposium was devoted to "The Structure and Function of Membranes" is a measure of the extraordinary advance made in that area over the past few years. Roughly the first half of the meeting was devoted to transmembrane proteins. These can be divided into two groups: those which extend across the bilayer in a simple manner, such as with a single α -helix—the "fibrous" proteins, and those which have an extensive mass in the plane of the membrane—the "globular" proteins (M. S. Bretscher, MRC Laboratory of Molecular Biology, Cambridge). Examples of the fibrous class include the erythrocyte MN glycoprotein and the spike proteins of Semliki Forest Virus. In the latter case, freeze fracture electron microscopy shows that there are no visible intermembrane particles in the fracture plane, (K. Simons, European Molecular Biological Laboratory, Heidelberg), which fits well with the view that these fibrous proteins have little mass in the plane of the membrane. All the other transmembrane proteins discussed were globular and in freeze fracture are visible as particles.

Had he been present, Walther Stoeckenius would have been pleased to see the progress made on two ordered membrane proteins which he first isolated—the purple protein (bacteriorhodopsin) and gap junctions. The purple protein is induced in a halophilic bacterium when growth conditions become unfavourable. Its assembly can be studied by addition of nicotine to the growth medium which inhibits the synthesis of the cofactor, retinal. Bacterio-opsin is formed which, on subsequent addition of retinal to the culture generates bacteriorhodopsin. In these conditions the bacteria have no purple membranes, however, just a brown membrane in which the bacteriorhodopsin probably exists as single molecules. Formation of the purple membrane seems to be more than just crystallisation of individual molecules of bacteriorhodopsin in the plane of the membrane, as the process requires energy (D. Oesterhelt, University of Wurzburg). The structure of the purple protein, as deduced largely from low

dose electron microscope image reconstruction, shows that little of the globular protein extends outside the lipid bilayer. About 70% of the protein can be accounted for by four α -helices which are perpendicular, and three which are tilted, to the plane of the membrane, (R. Henderson, MRC, Cambridge.) So far, X-ray studies of ordered arrays of gap junctions have been hampered by low short-range order. Nevertheless, the electron density profile suggests a rather even concentration of protein across the two bilayers constituting the junction, with about half the intercellular space filled in by protein. A tentative model for this array exists in which each pore is composed of twelve subunits: six in each bilayer (D. Caspar, Brandeis University). These two proteins, then, have taken on a defined shape: they no longer are drawn as wursts (objects with which the meeting became too well-acquainted at meals). Chemical characterisation of the gap junction protein is complicated by proteases, so that the subunit molecular weight of its constituent protein(s) is not yet known (N. Gilula, Rockefeller University).

X-ray or neutron diffraction studies of membranes can be aided by having an oriented specimen. Surprisingly, this can be achieved by applying a magnetic field: some membranes, such as erythrocyte ghosts, orientate

parallel to the field, others (purple membranes and rod outer segment disks) perpendicular to it. This may be useful, not only for orientating membranes, but for what it tells us about the overall polarity of α -helices in the membrane. Using this ordering technique, only slight changes occur in the X-ray diffraction pattern of rhodopsin in disk membranes on illumination—changes which are too small to be compatible with the model proposed by Blasie in which rhodopsin sinks a long way into the membrane on bleaching (M. Chabre, University of Grenoble).

A major function of specialised membranes is in energy production—pumping protons and using this gradient to make ATP, as proposed by Peter Mitchell. A beautiful system for studying the relation of proton gradients to electron transport and ATP synthesis is closed vesicles of higher plant chloroplasts, in which a gradient of 4 pH units can be generated by light. This pH gradient behaves as would be expected if it were an energy-storing device between electron transport and ATP formation. In addition, the pathway is reversible—artificially produced proton gradients drive reverse electron flow, which in turn induces light emission (M. Avron, Weizmann Institute).

The ion channel associated with acetylcholine receptors of denervated frog muscle can be studied with electrodes with exquisite sensitivity by limiting the surface area studied to a patch a few microns in diameter. The opening and closing of individual ion channels can be monitored and their properties, such as mean open time and mean conductance, can be determined. These results agree well with, and thus confirm, previous conclusions on the nature of ion channels obtained from membrane "noise analysis", and show that the current produced by a single channel has a square wave form (B. Sakmann, Max-Planck Institute, Göttingen).

The functional organisation of intracellular organelles is one of the most challenging areas of cell biology. A powerful approach to understanding the biochemical anatomy of higher cells is immunofluorescence microscopy, which has been used to special advantage to study the microfilament system in fibroblasts. This reveals that stress fibres contain, besides actin, myosin, tropomyosin and α -actinin. How the assembly of these structures is controlled, however, or why viral transformation (in SV3T3 cells) should lead to a gross diminution in the amount of fibres seen, is unknown. In addition, this technique has been used successfully to look at the disposition of microtubules and their polymerisation

Fv-1 function

IN the report of April 22, 1976 on the Armand Hammer Symposium at the Salk Institute (*News and Views*, 260, 669; 1976), it was indicated that our data on the function of the Fv-1 locus of the mouse showed that its product restricts growth of murine leukaemia viruses after integration of the provirus. At the time of the meeting, we had preliminary evidence in that direction which was discussed. More complete investigation, however, has shown that integration is prevented by the Fv-1 gene product although synthesis of proviral DNA is not prevented. Reports of our results and similar results of Sveda and Soeiro are in press in the *Proceedings of the National Academy of Sciences U.S.A.*

DAVID BALTIMORE
PAUL JOLICOEUR

in vivo after disassembly induced by cold or by colcemide. In the majority of cases, individual tubules are seen to grow in a unidirectional manner from the cytocentre of the cell towards the plasma membrane (K. Weber, Max Planck Institute, Göttingen).

Do any of these filamentous molecules interact with membranes? Possibly actin does with chromaffin granules, because when actin is polymerised at 30 °C in the presence of granules, the majority of actin fibres associated with granules show a specific polarity with respect to the granule. Staining with muscle heavy meromyosin shows the arrowheads on most fibres pointing away from the granule (B. Jockusch, Biozentrum, Basel). A different fascinating system in which actin is closely associated with a plasma membrane is found in starfish sperm. On reaching the egg, a 90 µm spike pops out of the sperm head, taking about 30 s to assemble. Artificial triggering, induced by ammonia, allows these spikes to be purified; they consist of an actin bundle surrounded by a limiting membrane. Before erection, this "profilactin" exists in a cup above the sperm nucleus. On isolation, the cup contains just three proteins—actin, and two proteins with molecular weights similar to human erythrocyte spectrin (250,000 and 230,000 daltons). How the spike is assembled and where all the surrounding membrane comes from, is unresolved (L. Tilney, Paris).

The interiors of higher cells abound with membranous structures. The only such component isolated in a pure state (apart from specialised secretory vesicles) is the coated vesicle. By a standard purification procedure, coated vesicles have been obtained from brain, adrenal medulla, parotid glands—all secreting cells—and from a lymphoma cell line, EL4, which is a non-secreter. The major protein in every case is clathrin (180,000 daltons) which seems, from one-dimensional fingerprints, to be highly conserved in sequence. This protein gives rise to polymorphic structures enclosing a vesicle: the outer diameter varies from about 500 Å to 1,200 Å. By tilting stained specimens in the electron microscope, the structures of some of the smaller coats have been deduced—they are made of twelve pentagons and a specific number of hexagons. The function of coated vesicles probably varies between cell types, but is believed to involve membrane transfer between organelles (B. Pearse, MRC, Cambridge).

Activation of cellular processes often results from an extracellular signal. In a simple activation, the time elapsed between signal and effect should be small and this is a virtue of the mast cell: when triggered by antigen, hista-

mine, which is prepackaged in vesicles, is released within seconds. Triggering requires that the antigen crosslink IgE receptors, but ferritin labelling and electron microscopy indicate that no extensive aggregation of receptors occurs in this brief time. Both energy and extracellular calcium ions are also required, and the latter must be present at the moment of antigen addition: later addition of calcium is ineffective. Calcium ions presumably enter the cell as a calcium ionophore alone will trigger mast cells. After triggering, exocytosis of histamine occurs. Ferritin labelling shows that, where the plasma and vesicle membranes become apposed, concanavalin A binding sites are excluded, as are particles seen in the freeze-fracture plane; however, the glycolipid GM1, which can be tagged with ferritin-conjugated cholera toxin, is not excluded from this area. Moments after exocytosis, the site of membrane fusion is confusion—lots of multivesicular blebs (M. Raff, University College London).

Blue light for aluminium nitride

from Andrew Holmes-Siedle

LIGHT-emitting diodes composed of electroluminescent materials are used mainly in numerical displays such as mileage indicators in cars, voltmeters and other similar instruments. Considerable interest was aroused in the years 1972–4 when blue and ultraviolet emission from gallium nitride electroluminescent devices was announced (see for example, J. I. Pankove, *J. Luminescence*, **1**, 114–126; 1973; *News and Views*, **252**, 443; 1975). It was a considerable feat of materials synthesis and device design by Pankove's group to make a structure which emitted fairly efficiently in the yellow region at a peak wavelength of 590 nm (corresponding to a photon energy of 2.1 eV) and also emitted some blue light and an ultraviolet line (378 nm, corresponding to 3.28 eV). This was a record low wavelength (that is, record high photon energy) for junction electroluminescence in the III-V compounds, some other members of which are now commercially established as light emitters at lower photon energies (GaAs, GaP, GaAl and their alloys). From the lack of commercial interest in gallium nitride devices since then, however, it seems that no-one particularly needs blue or ultraviolet-emitting diodes at the moment. Perhaps display

How do cells move? Cell locomotion and capping in lymphocytes are presumably related phenomena. If there exists an oriented lipid flow ($3 \mu\text{m min}^{-1}$) in the plasma membranes of motile cells, a membrane protein complex with a low diffusion coefficient would be capped and, if the complex anchored to a substratum, the cell would move forwards. If the lipid flow were much faster ($15 \mu\text{m min}^{-1}$), all proteins would cap, which might be unhealthy, whereas if the flow were rather low ($0.5 \mu\text{m min}^{-1}$), capping of surface antigens would not occur in lymphocytes. In other words, the rate of particle movement on a fibroblast ($3 \mu\text{m min}^{-1}$) is just right for the lipid flow hypothesis (Bretscher).

It seems to be that the overall problem of the structure of membranes is solved in outline, and specific aspects are well on the way to being solved in detail. But how membranes of a cell are inter-related, how they are organised by cytoplasmic components—that is where future interest lies. □

designers are happy with the green-emitting GaAsP alloys which the human eye responds to most efficiently and the red LEDs, made from GaAlAs, which are perhaps the most arresting. It is likely, however, that a use for blue emitters will emerge, especially in the field of data recording, where the extra photon energy could be needed to induce rapid photochemical effects. Thus, it is not surprising that IBM have attacked a new and formidable peak in the luminescent Himalayas of the III-V compounds, namely aluminium nitride (AlN). This compound is difficult to prepare in suitable form; it is refractory (melting point $>2,000^\circ\text{C}$) and is normally a good insulator, the resistivity of synthetic single crystals frequently being above 10^{10} ohm-cm . Since the principle of generating junction electroluminescence is to put a crystalline p-n structure under forward bias and pass several milliamps of current, these adverse electrical properties (which are quite consistent with the optical data indicating a forbidden gap of 6.2 eV) have to be somehow circumvented.

R. F. Rutz, reporting in *Applied Physics Letters*, **28** (7), 379; 1975), describes an empirical solution, interesting because it is, perhaps, the result of

Comparison of forbidden gaps, E_g , in III-V compounds

	Principal quantum number of valence electron	Atomic number	N	P	As	Sb
			2	3	4	5
		7	15	33	51	
B	2	5	5.5 eV	2.0 eV	1.45 eV	—
Al	3	13	6.2 eV	2.42 eV	2.17 eV	1.63 eV
Ga	4	31	3.5 eV	2.24 eV	1.43 eV	0.67 eV
In	5	49	< 1.9 eV	1.29 eV	0.33 eV	0.16 eV

As the energy gap becomes higher the emission moves further towards the ultraviolet.

some useful point-defect structure formed during an unusual type of crystal growth. A very thin crystalline film of AlN was first grown epitactically on a plate of single-crystal tungsten or sapphire by RF sputtering (this would, presumably have to be done very slowly and the crystal would have to be heated). This 'seed' film was then placed very close to a polycrystalline AlN disk in a reactive gas atmosphere and heated to about 1,850 °C, with the seed slightly cooler than the source. A single-crystal film of about 10 μm could then be grown and such films had an unusual colour, which suggested a high concentration of nitrogen vacancies. As a result of these vacancies, some of the films had quite a low resistivity (about 10^3 ohm-cm) and they appeared to possess n-type conduction. In this, Rutz's special growth techniques appear to have the advantage over that used earlier by RCA workers (Duffy *et al.*, *J. Electron. Mater.*, **2**, 359; 1973), who grew AlN films on sapphire but report no electrical measurements. The first electrical stimulation experiments by Rutz produced easily visible emission of blue light under the negative electrode, probably as a result of impact ionisation, with a peak in the ultra violet at 3.443 eV (360 nm) and quite strong emission up to 5.76 eV (215 nm), total output for one diode being in the region of a microwatt. The latter emission cutoff value handsomely exceeds the short-wavelength emission record set by Pankove using gallium nitride but quantum efficiency is low as yet, being in the 10^{-6} range.

The practical implications for data recording or display cannot really be assessed yet. The material is obviously more difficult to handle than the low-melting semiconductors, but, on the other hand, aluminium is much more abundant than gallium and the high photon energy opens the attractive prospect that very many phosphors can be excited by these emissions, thereby yielding any colour of display that we wish (this is, after all the principle used both in colour TV and fluorescent lamps). The felicitous use of vacancies to achieve a practically useful effect is also intriguing and will perhaps spur us to look further at the electrical stimulation of some other materials tradi-

tionally regarded as incorrigibly 'insulating' such as silica, alumina and—perhaps even the Everest of the III-V compounds—boron nitride. □

Ring chromosomes and DNA structure

from Martin Bobrow

THE development of relatively simple techniques for differentially staining sister chromatids in mammalian cells is an important addition to the armamentarium of the cytogeneticist. Latt first discovered (*Proc. natn. Acad. Sci. U.S.A.*, **70**, 3395; 1973) that BUdR-substituted chromatin is less intensely stained by the fluorescent dye Hoechst 33258 than are normal thymidine-containing regions. This can be exploited by allowing cells to replicate twice in the presence of BUdR; at the second metaphase, one chromatid will be fully substituted with BUdR, while the DNA of the other chromatid will have the original thymidine-containing strand and a newly synthesised BUdR-containing complementary strand. Differential staining of the two chromatids can be achieved using Latt's original fluorescent technique; by staining with Hoechst 33258 followed by Giemsa (Perry and Wolff, *Nature*, **251**, 156; 1974); with Giemsa alone after heat treatment (Korenberg and Freedlender, *Chromosoma*, **48**, 355; 1974), or by acridine orange staining (Kato, *Nature*, **249**, 552; 1974).

These techniques elegantly demonstrate the presence of sister chromatid exchanges. Such exchanges are a regular feature of normal cells. Their frequency is greatly increased in an inherited disease (Bloom's syndrome—Chaganti *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, **71**, 4508; 1974), and they are a sensitive indicator of certain classes of chromosome-damaging agents (see Savage, *News and Views*, **258**, 103; 1975). It is generally presumed that exchanges between sister chromatids reflect in some way a mechanism of DNA repair, and it is tempting (although still

very speculative to draw analogies with post-replication or recombinational repair mechanisms in bacteria.

An unusual and imaginative use of sister-chromatid staining techniques has been reported recently by Wolff, Lindley and Peacock (*Proc. natn. Acad. Sci. U.S.A.*, **73**, 877; 1976) to lend support to a controversial notion that there may be switches in polarity within DNA chains. Ring chromosomes which undergo a sister-chromatid exchange become a single continuous chromatin strand, since the exchange links each chromatid to its sister. At subsequent metaphases, such a chromosome appears as a dicentric, with its centromeres symmetrically placed, and a circumferential length twice that of the original ring chromosome. The principles of ring chromosome behaviour are those of the Mobius strip (well described by Gardiner, *Scientific American*, **219**, 112; 1968).

Wolff and his colleagues studied ring chromosomes induced by irradiating Chinese hamster cells in tissue culture. Conditions of culture were such that sister chromatid exchanges involved in the formation of the dicentric rings could be directly demonstrated. Cell division after irradiation was inhibited by colcemide, to prevent the dicentric chromosomes being disrupted on the mitotic spindle. The majority of the dicentric ring chromosomes with symmetrically placed centromeres did show the presence of a sister chromatid exchange as expected. A substantial minority, however, showed no such exchange present. After an exhaustive consideration of possible errors in scoring, and of the various mechanisms of ring chromosome formation, the authors conclude that such a ring could only be formed if the original chromosomal DNA (before ring induction) had contained a switch in polarity. The effect of a switch in polarity would be that when the broken ends of the irradiated chromosome rejoin to form a ring, the normal constraints of DNA structure would ensure that each chromatid end became continuous, not with its own tail, but with that of its sister. So in effect sister chromatid exchange and ring formation would be accomplished simultaneously, at the time of ring induction.

The authors believe that such reversals of polarity are an integral part of the structure of the chromosome, and estimate their frequency at one in 10^9 phosphodiester bonds. They could conveniently occur at positions of protein links in a DNA chain, but the existence of such links is in considerable doubt. The evidence today points towards each chromatid containing a single continuous DNA chain, in which case a reversal of polarity must imply the existence of 3'-3' and 5'-5' bonds.

It is obviously hazardous, in our present state of ignorance, to attempt to explain cytological phenomena at a molecular level. The explanation proposed does, however, fit the observation and no alternative explanation springs readily to mind. Whether the particular hypothesis is or is not eventually proven correct, the work of Wolff, Lindsley and Peacock is a fine example of how much of interest can be gleaned from intelligent microscopic observation of chromosome behaviour.

Facts of colonial life

from a Correspondent

A meeting on the Biology and Systematics of Colonial Organisms was held at the University of Durham, on April 7-8, 1976, organised by the Systematics Association.

MANY organisms live in colonies but the varied nature of coloniality suggests that such associations are influenced by an interplay of factors among which cooperation, competition for space and reproductive and feeding requirements are most significant. At the meeting the phenomenon of coloniality was traced from bacteria, through bryozoans and echinoderms, to sea birds in an attempt to discover what common factors link the phenomena of association, gregariousness and coloniality. Groups of animals of the same species may live in close proximity (gregarious colonies) and so form associations which differ from the integrated growing-together of "conventional" colonial organisms (corporate colonies).

The significance of cooperation was illustrated at microbial level (M. J. Carlile, Imperial College, London) for bacterial colonies. These aggregations are more able to withstand desiccation and show enhanced rates of enzyme production, improving substrate attack and serving as a barrier between colonies of different ages and genotypes. Cooperative coloniality occurs also in protozoans where, in more advanced taxa, some parts of colonies are already functionally specialised for gamete and spore production (C. R. Curds, British Museum (Natural History)). Proto-cooperation may be significant in gregarious colonies as in the bizarre and extinct, sedentary, rudist bivalves which maintained growth away from their substratum only with the support of contiguous neighbours (P. W. Skelton, University of Liverpool).

Higher levels of cooperation are

T4 chromosome

from a Correspondent

THE complex chromosomes of higher organisms are tightly organised, their DNA being coiled into standard "beads" which contain 200 base pairs associated with specific histone proteins. Recently the DNA of bacteria has also been shown to have a similar structure, although it is stabilised by a different kind of protein. Do bacteriophages, which are probably the simplest type of living organism, also replicate their DNA in a highly structured manner? In phage T4 it has been known for some time that intracellular phage DNA must be organised in some way, as the rate at which it sediments during centrifugation (about 700S) would reflect DNA concatemers with lengths of nearly 1,000 T4 genomes if attached end-to-end; but in fact the concatemers are at least 10 times smaller than this. The intracellular DNA must therefore be tightly folded and/or attached to some type of non-DNA molecule.

A surprising property of these structured T4 DNA molecules is their stability. Unlike the chromosomes of eukaryotes or bacteria their sedimentation rate is unaffected by treatments with protease, RNase, heat or detergent. One method of affecting this stability has been found however. When bacteria infected with a temperature-sensitive phage mutant in gene 32 (which controls the synthesis of unwinding protein) are grown at low temperature, normal intracellular DNA sedimenting at 700S can be isolated; but within one minute of shifting these infected cells to 42 °C

the phage DNA is found to have an appreciably lower sedimentation rate which is approximately that of T4 DNA which is now only one genome in length.

Curtis and Alberts (*J. molec. Biol.*, **102**, 793; 1976) have studied the reasons for this dramatic change in sedimentation behaviour in detail. They have shown that along the DNA concatemer which is formed during T4 replication there are sizable single-stranded regions spaced at average intervals of approximately one genome. Normally the product of gene 32 binds tightly to these single-stranded regions and protects them from nuclease action. This protection is rapidly lost if the 32 protein is inactivated, leading to strand scissions and the formation of phage DNA of single genome lengths.

On its own this new information concerning the existence and regular spacing of the single-stranded regions does not explain the higher sedimentation rate of normal intracellular T4 DNA. It does however suggest that single-stranded regions play a key role in holding the T4 DNA in a tight configuration, perhaps by attachment to some non-DNA substance in a way not too unlike the situation obtaining with the chromosomes of higher organisms and bacteria. As Curtis and Alberts point out, some high degree of order in the structure of DNA might be required to guide the simultaneous processes of DNA transcription, replication and recombination—and in the case of phage, of maturation as well.

seen in corporate colonies in which specialised polymorphs, or groups of polymorphs (cormidia) function at several hierarchically cooperative levels. This is spectacularly marked in bryozoans where an increasing number of colonial structures have been recognised as modified zooidal features common in colonies with very high levels of integration (J. S. Ryland, University College of Swansea). Colony-wide feeding and maintained current systems may be related to these levels of integration (P. L. Cook, British Museum (Natural History)).

By contrast, the level of cooperative coloniality in sponges is uncertain. Grafting experiments (A. S. G. Curtis, University of Glasgow) indicate that clones or clusters of sponges do not exhibit corporate coloniality. Grafts from a single sponge body will be accepted by another only if they are of the same strain. Similar strains will

coalesce if they grow adjacent to one another. Although it was argued (W. G. Fry, Luton College of Technology) that a sponge is a colony in which aquiferous units represent individuals it is evidently difficult to define the individual in sponge morphology and the implication of Curtis's experiments were of wide interest to participants at the meeting.

Reproduction as a factor in coloniality is significant in that most corporate colonies develop by asexual budding. The factors controlling the timing of the appearance of new individuals are not known. Indeed no obvious relationship to age or maturity for asexual budding can be deduced in, for example, rugose corals. Asexual reproduction in corporate coloniality links closely with the cooperative function but sexual reproduction may itself be a cause of aggregation in some organisms. Barnacle aggregations relate to their

internal fertilisation requirements since they have to settle sufficiently close together for this to occur (D. J. Crisp, University College of North Wales). Although external fertilisation occurs in the sessile marine gastropod group of the vermetids proximity seems equally necessary for their survival (R. N. Hughes, University College of North Wales). Reproduction is not, however, so basic a factor as environmental conditions in the spacing of serpulid worms since the same species may occur singly as well as in aggregations (H. A. Ten Hove, State University of Utrecht). This is also shown by clustering of serpulids in Ardara Lough, County Galway, where they aggregate densely on restricted areas of suitable substrate in otherwise favourable parts of the Lough (D. W. J. Bosence, Goldsmiths' College, London).

The final paper of the meeting (J. C. Coulson and F. Dixon, University of Durham) linked coloniality and sexual reproduction in a study of kittiwake colonies in which the protection and stability afforded by colonial association supports successful and constant breeding pairs in inner areas whereas in outer areas of colonies breeding and adult survival rates are lower and divorce rates higher. Even so the system is not inbreeding because young birds are recruited from other colonies.

The role of suspension feeding in coloniality is important for echinoderms (G. F. Warner, University of Reading) and invites generalisation for other groups also. Suspension feeding allows crowding of aquatic benthonic organisms but, as in the case of starfish, dense aggregates may be a predatory response to food shortage.

Competition for space was mentioned by many contributors but discussed in its own right by J. B. C. Jackson (The Johns Hopkins University). In an analysis of the relative success of species on hard substrata of tropical reef environments the particular success of corporate colonial organisms in two aspects essential for establishment was emphasised—effectiveness in competing for space (outward growth) and effectiveness in establishing a particular level in the water column for food supply (upward growth). Thus colonials always prevail over solitaries even though colonials generally settle later. Solitary sedentary species hold their own if they aggregate in the ways described for barnacles, vermetids, rudists and serpulids, or if they grow upward.

For the four main themes of the meeting the starting point is cooperation. In marine organisms this is made possible (necessary for some) by the suspension feeding habit. Sedentary marine organisms must also compete for space. Alternatively competition for

space also leads to gregarious coloniality which arises when sexual reproduction rather than cooperation alone may be significant. Such generalisations derive from consideration of the groups referred to at the meeting. These also included coelenterates, ascidians and graptolites but not social insects, fish schooling behaviour or gregarious and social behaviour in mammals. Perhaps cooperation, space competition, feeding and reproduction are also consistent themes which can be carried through to the most complex colonial/social species of all, *Homo sapiens*. □

Informatics

from Kevin P. Jones

Informatics 4, organised by the Aslib Co-ordinate Indexing Group, was held on April 5–7, 1976 at the University of Lancaster. The proceedings will be published by Aslib.

THE original Russian definition of informatics, by Mikhailov, Chernyi and Giliarevskii (*Osnovy Informatiki*, Moscow, 1968), that it is the learned discipline which studies the structure and properties (but not the actual content) of scholarly-information activity, its theory, history, methodology and organisation, provided the starting point for several papers. If this definition is taken with R. A. Fairthorne's assertion (*Informatics 3*, Aslib, in the press) that the target for informatics should be the automation of skills based on linguistic experience, together with the fact that the conferences have been deliberately interdisciplinary, then it should be possible to gain some impression of what the *Informatics* conferences have set out to achieve.

The inter-relationships between linguistic structure and meaning have provided a dominant theme through all the conferences. These may be traced in Margaret Masterman's (Cambridge Language Research Unit) contributions, which have developed the hypothesis that as it is now technologically feasible to store what amounts to a 'library' on magnetic disks then it is desirable to develop techniques for extracting information from them while retaining its original meaning. Her theory has explored the automatic partitioning of text by the use of rhetorical figures and the applicability of LIST processing languages, notably LISP 1.5, to semantic structures.

The analysis of structure within large textual units was explored by W. John Hutchins (University of East Anglia). Clearly the elucidation of this structure

must form the basis for decisions on what a text is about in both manual and automated systems. Interestingly, this paper, which had been developed from the standpoint of indexing need, shared common ground with Yorick Wilks's (University of Edinburgh) exploration of frame theory from artificial intelligence: both commented on story structures. Wilks was critical of frame structures, particularly those which attempt to map the system's (or robot's) environment as the structures tend to alternate between inadequacy and over-complexity. It is clear that it is going to be extremely difficult to mirror by machine man's ability to focus on his environment and correlate this with earlier experience.

Two lively interactive systems were described. One was developed at the University of Aston by John Fisher (Herriott Watt University) as part of his development of a *Functional Theory of Language*. Verbs, notably an explicitly performative class, achieve an unusual prominence in Fisher's theory and careful categorisation of these has enabled highly inferential programs to be written which are capable of generating a genuine man-machine dialogue. Robert Oddy's (University of Aston) novel approach to information retrieval bridged the narrowing gap between artificial intelligence and information retrieval. Oddy's program replaces normal Boolean search strategies by an automatic modelling technique which attempts to reflect the enquirer's needs by constructing large semantic networks.

The history of the thesaurus, or conceptual wordbook, was reviewed by Tom McArthur (Longmans), from Bishop Wilkins through Roget to the author's *Lexicon of Contemporary English*. Thesauri have been widely exploited in the design of information retrieval systems to provide semantic frameworks for assisting indexing and searching operations.

Informatics originated in the USSR and it was pleasant that Professor Vladuts was present at the conference, from VINITI—his formal contribution consisted of a survey of current Soviet research within information languages.

Fairthorne, who has been closely associated with Informatics in the United Kingdom, has frequently castigated librarians and information scientists for their disregard for theory. Informatics 4 not only showed that a broad base for theoretical development is fully justified, but that the careful application of theory can lead to great improvements in practice. Peter Noerr (British Library) has applied computers to a bibliographical classification using relational operators, an operation thought to be impossible only a few years ago. □

reviews: telecommunications

Trends in optical fibre transmission research

J. E. Midwinter*

Recent advances in the manufacturing technology of low loss optical fibres have made possible serious systems design and testing. But this has highlighted the fact that the mechanisms involved in the propagation of light in multimode fibres were at best, poorly understood. New measurement techniques are now being used to check more sophisticated theoretical analyses of propagation in "imperfect" and cabled fibres and are leading to a deeper understanding of this new transmission medium, its possibilities and its limitations.

THE invention of the laser in 1961 and particularly the GaAs laser in 1962 led research workers in the telecommunications field to examine the possibilities of using light in place of electricity for signal transmission. Free space propagation could be ruled out for civil use because of the relatively frequent occurrence of fog, smoke or rain causing interruptions, and at a fairly early stage, glass fibres were proposed as a suitable medium for guiding the light from source to detector. Stimulated, no doubt, by the observation that the carrier frequency of light is very large (10^{16} Hz), the early system proposals tended to be centred on very wide band transmission, typically 10^8 to 10^{10} Bit s^{-1} combining single mode fibres and lasers to form a sophisticated transmission system. Now, a decade later, real systems are beginning to emerge from the laboratories and surprisingly to the outsider, the intervening research has led to less exotic specifications with most systems designed to operate in the range 8–140 MBit s^{-1} using very overmoded fibre and in many cases with an incoherent light-emitting diode source in place of a laser. But at the same time, this apparent down-rating of the performance has been accompanied by an increasingly more detailed understanding of the transmission medium, its limitations and possibilities.

The shift to lower data rates results from a careful appraisal of the real needs of the telecommunications networks and also from a rather more sober evaluation of the capability of complete systems as compared to the theoretically best performances attainable from individual components. The GaAs laser remains a most attractive source, both on grounds of ease of modulation, through its drive current, and because of its near ideal wavelength coinciding with a minimum in the loss spectrum of the fibre. Interest also centres around the use of light-emitting diodes (LEDs) which currently offer rather longer life expectancy and sufficient performance when set alongside lower bandwidth requirements and more sophisticated fibre designs.

An LED source requires the use of a large core diameter (60 μm) fibre if sufficient power is to be launched, while it is much easier to launch a laser efficiently into a large core fibre than into a single mode (3 μm core diameter) fibre. Finally, the need to join fibres in a field environment adds further to the attraction of using large core diameter fibres for operating

systems, although maximum performance will only be obtained using the very low dispersion single mode fibres. At present, however, a great deal of work is being devoted to understanding in depth the multimode fibres, waveguides which allow several thousands of electromagnetic modes to propagate with the energy spread nearly uniformly among most of them. This is to be contrasted to most microwave waveguide systems in which either the waveguide is designed to propagate only a single electromagnetic mode, or the launching and guide design are such as to favour strongly one mode only (as in the TE_{01} millimetre waveguide used for telecommunications). It should also be noted that in optical systems, one is concerned with waveguide propagation over distances of the order of 10 km using a wavelength of 10^{-4} cm; that is, over a distance of the order of 10^{10} wavelengths. Since the equivalent guide length in a centimetre waveguide would be 100,000 km we can expect to find a different emphasis in our study of the optical guides.

The early studies of the very overmoded optical fibre guides used a simple ray analysis to study the details of propagation but more recently more sophisticated approaches have been developed which still avoid the frightening complexity of a full electromagnetic wave analysis but which allow study of details such as the effects of different types of mode coupling and of core-cladding refractive index profiles which do not involve simply a sharp interface but which include diffusion or controlled rounding. Very recently, this analytical advance has been supported by experimental advances which promise to make possible the intimate study of overmoded guides, mode by mode, even when several thousand modes are present.

Before discussing these matters, however, we should note that this interest has only arisen since it has become clear that optical fibres will be available with very low losses and to reasonable tolerances (a few % in most vital parameters). Fibre losses in the range 2–10 dB km^{-1} have been produced by a number of distinct processes. Systems designs begin to be most attractive once fibre losses drop below about 7 dB km^{-1} .

Three fibre manufacturing processes seem to be in contention. The Corning Glass Company have described the "soot" process in which $SiCl_4$ with or without some dopant is reacted with oxygen in a flame to generate a stream of ultrafine and ultrapure silica particles which are deposited on a suitable

*Post Office Research Centre, Martlesham Heath, Ipswich, Suffolk.

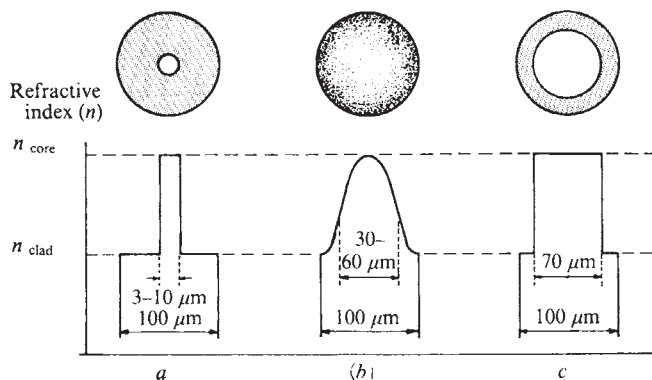
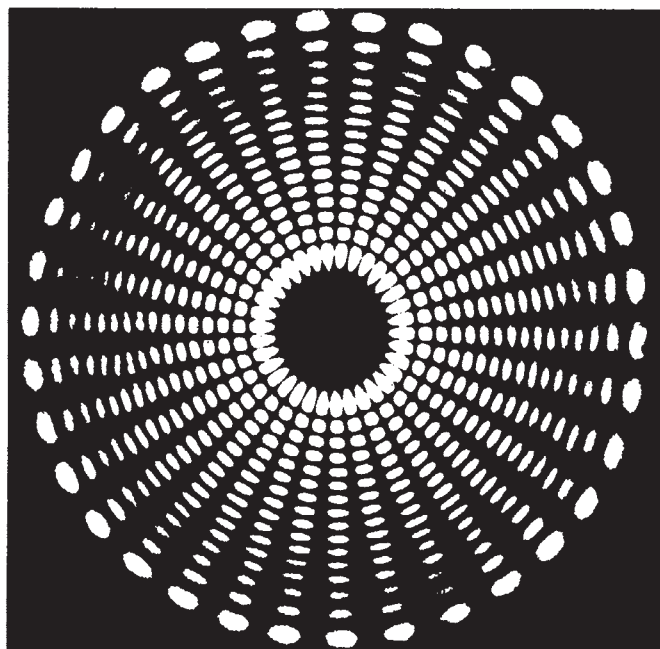


Fig. 1 Cross sections of the major fibre types showing (a) small core monomode fibre, (b) large core step-index multimode and (c) large core graded-index fibre. In all cases the index difference between core and cladding is about 1%.

mandrel. Two thick layers of doped silica are built up and sintered, the mandrel is removed and the resultant mass collapsed to form a preform having an inner core of higher index material surrounded by lower index material with differing dopant concentration. This is then pulled in very controlled conditions to yield fibre. The second process is very similar, except that the gases are reacted in a hot zone within a silica tube, so that layers are built up on the inside surfaces of the tube. The tube is then collapsed to form a preform and pulled to yield fibre. Both processes have the attraction of working directly from the vapour phase to yield very high purity solid material. This is vital because the principal loss mechanisms in glasses in the 800–900 nm wavelength band are absorption by iron, copper, nickel and manganese contamination.

The remaining process relies on glass made from oxide powders, produced in more or less conventional conditions in bulk. It is then converted into fibre using two concentric glass streams from which a fibre can be drawn. In this process exceptional care must be taken not to allow any contamination to enter the pure starting powders used, and chemical means must be used to maintain the impurities that are present in the most favourable oxidation state if their absorption loss is not to be unacceptably high. The process does have the attraction that it is immediately usable for continuous production, since glass can be loaded into the crucibles at the same

Fig. 2 The mode pattern of an $LP_{17,16}$ mode (kindly supplied by W. Stewart of Plessey Co.).



time that fibre is being drawn, whereas preform processes are run in batches and are frequently limited to generating lengths of only a few kilometres of fibre per preform.

Given low loss fibre, which is essential if a fibre communications system is to compete with existing coaxial cable systems, then the fibre must be engineered into a system in such a way as to allow useful amounts of information to be transmitted over the long distances made possible by the low loss, typically 5–10 km of fibre between source and detector. Thus the dispersion of the fibre becomes significant.

If one focuses attention initially on a step index, multimode fibre as shown in Fig. 1c then one can imagine light propagating along the higher refractive index core structure by total internal reflection at the core-cladding interface. Since it is a large core diameter fibre, the energy can travel in any one of several modes, zig-zagging at a large angle or a small angle to the axis, or following various helical paths. These modes or paths are of different length and suffer different time delays. Since in practice, most will be excited, very large pulse dispersion occurs, so limiting the available bandwidth. For all modes excited equally, the r.m.s. pulse width after length L is given by

$$\tau = \frac{L\Delta}{2\sqrt{3}c} \quad \Delta = (n_{\text{core}} - n_{\text{clad}})/n_{\text{core}} \quad (1)$$

$N = \text{group index of core}$

This amounts to 15 ns km⁻¹ for a value of $\Delta = 0.01$ which is typical. Such dispersion over an 8 km link would limit the 3 dB bandwidth to about 4 MHz.

Measurements of long lengths of step-index fibre show larger bandwidths than would be predicted by the above result. This arises from two factors, namely partial excitation of the available mode spectrum and mode coupling. If the source only excites the lower order (low angle) modes of the fibre, then the slowest modes will carry no energy and less pulse dispersion will result. Impressive bandwidths have been reported in such conditions but they are probably not of much more than academic interest since a fibre packaged into cable will in all probability be subject to mode coupling from the impressed lay deformations. This leads to subtle effects.

In the first instance, the launched energy spreads out among the available modes, populating both the slow travelling and fast travelling, so that compared with the earlier example, an increase in pulse dispersion will occur. But if the fibre was initially excited by an incoherent source such as an LED, most or all modes will have been populated from the start, so that the possibility of a low initial dispersion does not arise. Then the immediate effect of mode coupling is to start interchanging energy among modes. In this situation, after a certain distance, it can be said that any particular packet of optical energy will have spent some portion of its time in every mode, so that its mean group velocity starts to approach an average value, which is independent of the mode launched.

To describe theoretically such a state of propagation, it is necessary to make some assumptions about the nature of the mode coupling between the guided modes and the loss mechanisms that will be active within the fibre. Estimates can then be made of the pulse dispersion in such a situation. At this point some experimental guidance is required but before examining the data, we should refine our concepts of modes in the circular dielectric guide.

We have already commented that energy can propagate along zig-zag or helical paths in the fibre, with successive reflections at the core-cladding interface giving the guidance. Interference restricts the angles at which propagation can occur and so defines the "modes". The modes can best be described using the linearly polarised (LP) notation. An LP_{mn} mode is easily recognised as one having n field maxima along the radius vector and $2m$ field maxima around a diameter. The modes are all circularly symmetrical. Figure 2 shows an LP mode field pattern photographed on the end of a 60-μm core diameter fibre, with the core-cladding interface corresponding roughly

to the outer edge of the field pattern. The integer m is a measure of the degree of helical propagation, the greater m the tighter the helix, whereas the integer n relates to the angle the waves make to the fibre axis in the r - z plane. Any particular LP_{mn} mode in a given fibre can be characterised by a precisely defined value of propagation constant along the fibre axis or z direction, signified by k_z . Modes for which $(2n+m)$ are equal have roughly the same value of k_z , a fact which leads to the frequent use of a reduced mode number $M = (2n+m)$.

A study of the experimental data on mode coupling in fibres leads to the conclusion that the effects observed can best be described by couplings between modes having $\Delta M = \pm 1$ or nearest neighbour modes. Theoretical studies in which the likely mechanisms to induce such coupling have been considered suggest selection rules of the type $(\Delta m = \pm 1, \Delta n = 0)$ or $(\Delta m = \pm 1, \Delta n = \pm 1)$. Evidently the earlier conclusion includes the latter but is rather broader. To produce coupling of the latter type, all that is necessary is a small deviation of the fibre axis from the straight path. The deviation takes the form $\Delta x = d \cos(\Delta k_z z)$ where z is the direction of the fibre axis. For strong coupling to occur between modes having values of k_z given by $k_{z,1}$ and $k_{z,2}$ the relationship

$$\Delta k = |k_{z,1} - k_{z,2}|$$

must hold. In addition, the values of m and n for the mode pair must satisfy the selection rules above. Such restrictions sound sweeping, so that one might imagine that such couplings would be unlikely to occur, until it is noted that the values of Δk that are operative are typically 10 – 100 rad cm^{-1} . Therefore wavelengths in the millimetre to centimetre range are important while d can be less than 10^{-4} cm ! Such deviations can easily occur in a fibre packaging process, or even in winding a fibre on a drum so that we find in practice that such mode coupling is often present in fibres.

Evidently, coupling between different mode pairs will require different values of Δk_z . It is found that to couple modes with low values of M , smaller Δk_z values are required than for larger M values. This fact opens some possibilities and brings us to some problems¹. Figure 3 shows the allowed values of m and n for guidance to occur with a given fibre. Modes exist outside the boundary but they are lossy. Alternatively, one can say that the rays corresponding to the m, n values outside the guided-mode boundary are not totally internally reflected at the interface. If the mode coupling occurs in the fibre because of randomly distributed microbends, as deduced earlier, then we must assume that not only will guided modes be coupled to guided modes, but guided modes will also be coupled to leaking modes, resulting in an increased loss for the fibre. Experimentally this is found to be the case and the observations fit quite well with theoretical predictions.

Mode coupling is expected to reduce the pulse dispersion after an initial fibre length in which the first group velocity averaging takes place. If the additional loss due to the mode coupling is $\delta \text{ dB km}^{-1}$, then the initial distance for averaging is roughly given by $L_c = 1/\delta$. Up to that point, the pulse dispersion is described by the formula of equation 1. Afterwards, the pulse dispersion is reduced by a factor R where

$$R \approx (1/\delta L)^{\frac{1}{2}} \quad (2)$$

so that from this point on, the pulse width only increases as $\sqrt{L}$ instead of L . The effect of this can be very significant on long fibre lengths, as is shown in Fig. 4, in which the pulse dispersion is shown for a 1% index difference fibre with varying degrees of random mode coupling introduced. The curves are theoretical, but good general agreement with experiment is found. The almost inevitable presence of a small level of mode coupling in a cabled fibre thus leads to a considerable uncertainty as to how it will perform when in a very long length, such as is expected to be used in operating systems. This inevitably raises questions about the stability of such mode coupling

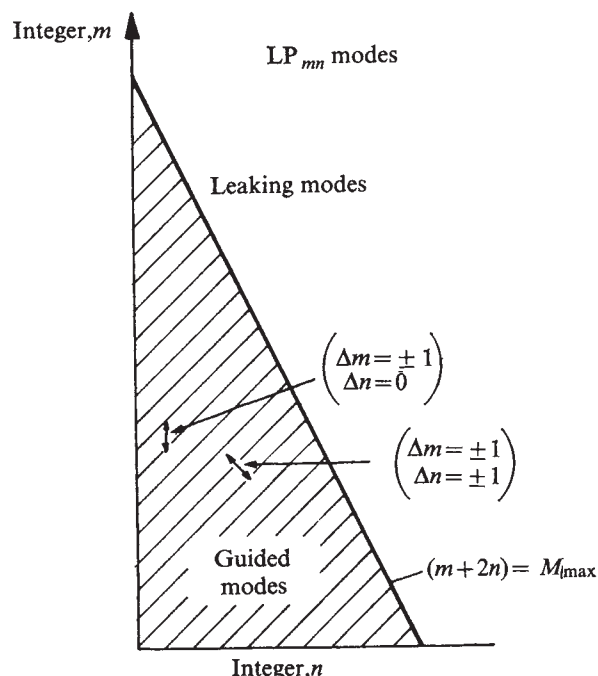
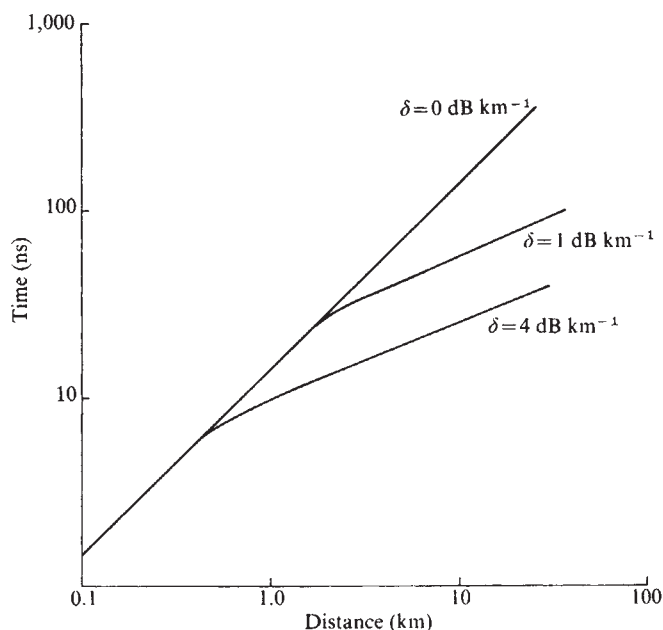


Fig. 3 The allowed values of m and n for the guided modes of a multimode fibre, showing the selection rules for coupling due to fibre microbends.

with time and the extent to which it can be exploited to enhance the fibre performance. Answers are still being sought.

We alluded to another interesting possibility that arises from this line of thought. It should be possible to introduce selective coupling only between guided modes of a fibre if the deformation spectrum (Δk) can be controlled. It is at least theoretically possible to introduce desirable amounts of mode coupling without loss but with improved bandwidth. To do this, the higher frequencies present in the spatial deformation must be cut off just short of the values at which coupling across the guided mode-leaking mode boundary occurs. Experimentally this has not been done, although experiments have been reported in which chosen pairs of modes have been selectively coupled together by the use of an impressed deformation. It remains to be seen how far the technique can be extended.

Fig. 4 The pulse broadening predicted by equation 1, both with and without small levels of mode coupling present. Calculated $\Delta = 0.01$.



A quite independent approach to achieving increased bandwidths with large core diameter fibres has been to develop fibres in which the core cladding interface is not sharp but is diffused or graded, as illustrated in Fig. 1b. The effect of such a profile rounding is readily understood, if we consider the ray paths for the three primary modes of propagation. A central ray travels in the core centre, in high index material by the shortest possible path but relatively slowly. The meridional ray, crossing the central axis, is focused back and forth along a near sinusoidal path, partly in low index and partly in high index material along an intermediate distance path but with a mean velocity given by some average of the indices traversed. Finally, the helical path ray travels the longest physical path, but wholly in low index material and thus at the highest mean velocity. Evidently, the path length divided by the mean velocity is to some extent compensating for each. Mathematical analysis, using the WKB^{2,3} method shows that a very high degree of compensation is possible provided that the index profile can be sufficiently well controlled. The form of profile that has received the most attention, termed the "alpha" profile, is as follows

$$n(r) = n_0 [1 - 2\Delta (r/a)^\alpha]^{1/2} \quad a > r > 0 \quad (3)$$

$$n(r) = n_0 (1 - 2\Delta)^{1/2} \simeq n_0 (1 - \Delta) \quad r \geq a \quad (4)$$

It is an attractive one to study theoretically, not only because analytical results can be obtained for it but also because it encompasses a range of profiles of experimental interest. For example, $\alpha = \infty$ corresponds to the step-index fibre discussed earlier, whereas $\alpha = 2$ corresponds to the parabolic dielectric constant profile that was the subject of much earlier study.

The outcome of these studies has been to show that the optimum profile is one that is very close to parabolic ($\alpha = 2$). The pulse dispersion in such a profile with all modes excited is decreased over the step-index profile by a factor of approximately $(4/\Delta)$ so that for a 1% index difference fibre, an approximately 400 times decrease in dispersion is possible in theory. Experimental results approaching this have been reported. An impressively low figure of 0.18 ns km^{-1} has been measured on a 1 km length of fibre, caused by mode dispersion. This represents an improvement factor of some 80 times over the step-index equivalent fibre. To obtain such performance requires, however, exceptional control of the fibre preparation stages and it has yet to be demonstrated that that is possible in large scale production. Against this, it must be said that many systems currently envisaged do not require such low dispersion in a fibre. It follows that fibres with carefully controlled index profile are of great interest at present, with many groups involved in the study of their detailed propagation characteristics, measurement of their vital parameters such as the actual profile obtained, together with intensive studies of manufacturing techniques for such fibres. The gains that can be achieved are illustrated in Fig. 5 which shows the theoretical dispersion curves for fibre with fixed $\Delta = 0.01$ but different values of α . Relatively little profile rounding greatly decreases the mode dispersion.

It is important to note one other factor that can seriously limit the performance of a fibre in reality, namely material dispersion. The figure quoted above of 0.18 ns km^{-1} for the mode dispersion in the fibre was not the measured total dispersion. The measured result was 0.25 ns km^{-1} , the difference being accounted for by the presence of dispersion from the material of the fibre itself, the dispersion that is used to good effect in prism spectrographs. The amount of pulse broadening from this effect depends on the linewidth of the source used. In the above case, a broadening of about 0.18 ns km^{-1} was due to this effect, corresponding to the use of source whose linewidth was approximately 1 nm. The dotted curves in Fig. 5 show the dispersion resulting when a source linewidth of 2 nm is used.

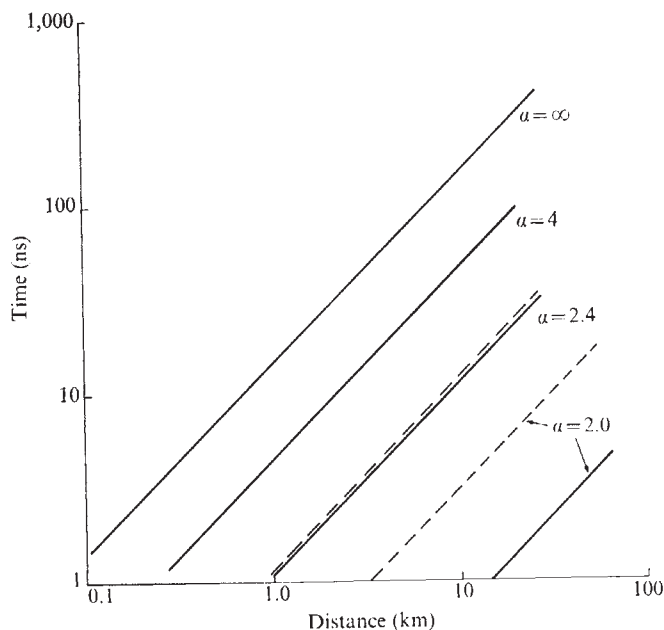
The presence of material dispersion thus sets a limitation

on the choice of source, its linewidth being of paramount interest. For example a GaAs laser source with linewidth of 2 nm gives rise to material dispersion of approximately 0.3 ns km^{-1} whereas a GaAs LED with much greater linewidth, say 20 nm, generates material dispersion of 3 ns km^{-1} , limiting its use to only the lowest data rate systems. An interesting alternative lies, however, in an observation first made by Kapron and Keck that the material dispersion (governed by $d^2n/d\lambda^2$) passes through zero for silica, the most popular fibre material, in the region of $1.2\text{--}1.3 \mu\text{m}$. This also corresponds to a favourable low loss transmission region in most fibres, but did not initially attract any interest because of the absence of both sources and detectors to operate there. Recent work on GaInAs LED sources suggests, however, that some reappraisal of the possibilities may be due, since operation in this spectral region would apparently make possible wide band optical systems using large core graded-index fibres with LED sources, thereby side-stepping many of the problems inherent in developing laser sources.

It is evident from the discussion of propagation in optical fibres that the design of cable for optical systems is likely to present some new problems. The most immediately obvious arises from the requirement that the fibre should not be subjected to microbending of the type likely to induce unwanted mode coupling. This means in practice that the fibre must be well cushioned within a small longitudinally rigid structure. An attractive theoretical design would be a rigid walled tube, filled with an easily compressed material in which the fibre was embedded, so that its natural stiffness allows it to lie straight within the spongy environment. Practical packaging to date involves extruded plastics chosen to balance cushioning, resistance to exterior forces and stiffness, in a single extruded coating, with nylon as a popular material. An alternative approach that is nearer to the theoretical optimum is the use of a polymer tube, in which the polymer molecules have been longitudinally aligned by stretching the tube after extrusion and within which the fibre lies loosely. The molecular alignment offers a tenfold increase in longitudinal Young's Modulus, generating a relatively very stiff tube while the loose fibre is free to flow over any irregularities that remain. Both approaches have led to reports of packaged fibre with no measurable ($\pm 0.5 \text{ dB km}^{-1}$) change in loss from microbending caused by the packaging process.

Once the fibre has been successfully cushioned against its

Fig. 5 Theoretical dispersion curves for graded core fibres with different values of the parameter α . The dotted curves include material dispersion characteristic of a 2-nm linewidth source.



immediate environment, it must be enclosed into a cable designed to allow continuous lengths of order 1 km length to be pulled into an underground duct. This raises problems which take one into a new realm of glass technology. A sizeable amount of information is available on the strength of glass fibres. It shows that small diameter or short fibres are stronger than large or long ones, in terms the mean strain permissible before breakage occurs. A great deal of study has been devoted to the development of the theory of microcracks to explain the observed behaviour of glass samples. The characteristic of glass, well known to everyone, is that it deforms elastically to a certain strain level and then fails catastrophically. There is no plastic flow region as there is with metals. The strain that a given sample will take before failing depends critically on its surface state, since small or invisible cracks or defects in the surface lead to premature failure because of the stress concentration that occurs around them. The result is twofold. The surface of the sample must be protected as thoroughly as possible and there are a number of proprietary techniques for doing this. But recognising that no protection is likely to be perfect, it must be assumed that the effect of such protection will only be to reduce the probability of failure in any given length. Laboratory based measurements have been confined almost entirely to fibre samples of the order of 1 m long. Suddenly one is concerned with the failure statistics of fibre lengths that are longer by a factor of 10^4 .

Extrapolation of the present data by four orders of magnitude is clearly likely to lead to misleading conclusions, so that experiments to test long lengths of fibre are being attempted, primarily aimed at verifying the existing statistical model developed by Weibull to describe such effects. A few tests of the failure probability for a given load applied for a short time have been made now on multi-kilometre lengths and the results shown to agree with the general predictions of Weibull. These were made by reeling long lengths of fibre from one reel to another under progressively increasing tension until breakages occurred.

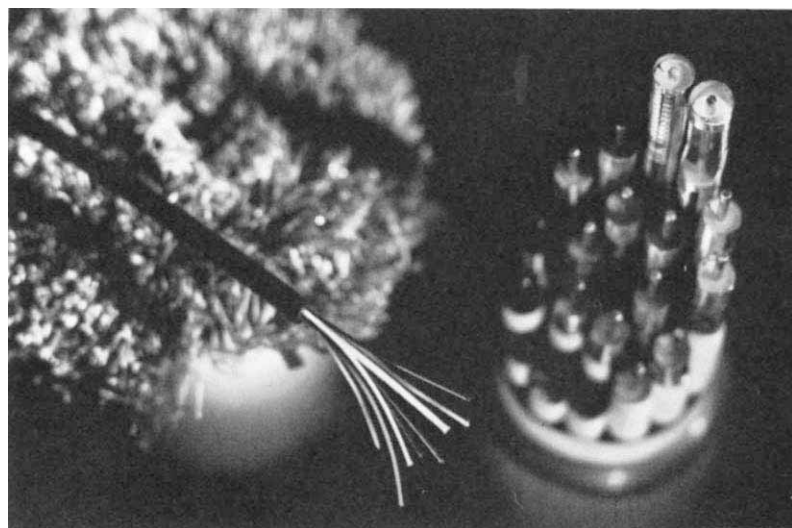
A second factor that requires further study is, however, the effect of microcrack growth in a glass filament when under a small but continuous strain. It is known that a fibre under a maintained strain well below its immediate failure level, will fail eventually because of the gradual growth of small cracks. Obviously, a fibre in a cable which has been pulled into a duct is likely to remain in a state of strain after the cable has been installed. At present it is not known what residual strain will be acceptable in the fibre to ensure an acceptably low failure rate over 10 or 20 years, although the initial indications suggest that strains of as low as 0.1% are the maximum allowable. To achieve such low levels of strain in cabled fibre, assuming that lengths of 0.5–1 km are to be pulled into ducts at a time, presents a new challenge to the cable designer. For comparison, we note that many present telecommunications cables can suffer strains of up to 10% in extreme cases without any major failure or serious degradation in their performance.

Having discussed some of the possibilities of this new transmission medium we shall now turn to the question of measurement techniques used to study and characterise it. It presents to the system engineer some interesting new problems not previously met in the study of existing transmission media. The most general point is that it presents in an acute form the problem that what you measure largely depends on how you measure it. To illustrate this cryptic comment, we refer back to the preceding discussion and consider how one should measure the pulse dispersion or bandwidth of a fibre transmission line and we compare this to a comparable measurement for a coaxial line. In the latter case, a fast (δ -function) pulse would be launched into the line at $L = 0$ from an impedance matched source and at distance $L = D$, a matched detector would record the received electrical signal which could then be analysed. Interpretation of the result might be complex, but the experimental result is unlikely to be very sensitive to the measurement situation.

In the case of the optical fibre, we can postulate a large number of possible measurements. We might take a short fibre length, and a single mode laser, launch a fast pulse into a single guided mode and detect the result before it had travelled far enough for significant amounts of energy to be coupled into other modes. We would have measured the very small dispersion associated with a single mode of the fibre. Or again, we might take a large area diffuse LED source, illuminate the whole end of the fibre exciting all modes, and again after a length short enough for there to be no mode mixing, we measure the pulse broadening. From this we can again deduce pulse dispersion, but now it will be a combination of mode dispersion and material dispersion, most probably dominated by the former. In either of the two former situations, we could allow propagation over a longer length so that mode coupling and mode dependent loss effects begin to play an effect. Ultimately a nearly stable mode distribution would be obtained, a distribution of energy among modes that once established will remain essentially constant. We can now remeasure the dispersion associated with the fibre. It will follow a curve of dispersion versus length of the general form shown in Fig. 4 with a linear initial growth, followed by a square root growth thereafter. Evidently, this is a much better measurement, provided we can clearly separate the linear and square-root regions.

Assuming this to be done in the laboratory, however, nagging questions remain. The last result depends critically on the amount of mode coupling present in the fibre. If the bare fibre is stored on a drum, this is sensitively dependent on the fibre tension, which controls how closely it follows microbumps on the reel surface. In a cable, similar considerations are likely to apply although the magnitude of any effect will evidently depend very much on the precise features of the cable design. We should also note that it is not only pulse dispersion that depends on the mode coupling and modal power distribution, but also the other critical design parameter, power loss against length. There is a unique loss rate (in dB km⁻¹) associated with a given stable mode distribution, but the region leading to that distribution may feature a higher or lower figure depending upon the launching characteristics. Both figures will change if the degree of mode coupling changes. In summary, we can say that there is not a single figure that can be quoted for either loss or pulse dispersion to characterise any particular multimode fibre. To understand the propagation in the fibre, a much more sophisticated model must be adopted, but out of this, guide rules will have to be distilled that can be used by the systems engineer, rules that will allow him to design a given link with confidence that sufficient margins will have been incorporated. In the short term, then, a great challenge is presented to optical measurement technology to thoroughly elucidate the details of transmission in the deceptively simple looking, yet very complex optical transmission line. A number of approaches are being developed.

It is evident from the above discussion that a measurement process must at least allow one to study propagation on very long lengths fibre, comparable to the lengths expected to be used in an operating system (say 10 km). But in addition, if the relative parts contributing to the overall dispersion and loss are to be resolved to allow design predictions to be made, measurements are necessary on intermediate lengths. The one technique that comes closest to achieving these requirements in a single step is the shuttle-pulse dispersion measurement, which is very reminiscent of echo measurements made on imperfectly terminated transmission lines. Experimentally a length of fibre, of perhaps 500 m or 1 km total length is mounted between high reflectivity mirrors with a source launching into one end and a detector at the other end. The launched pulse shuttles between mirrors, and the leakage power after every two transits is measured on the detector. By careful design of the apparatus and with low loss fibre, multiple transits can be observed, allowing a plot of dispersion against integer multiples of length to be made. Experimental curves



An optical fibre cable with, for comparisons of size, weight and raw material consumption, a 4,800-pair distribution cable (which carries local telephone traffic from junctions to individual receivers—maximum capacity 4,800 two-way circuits) and 18-core 9 mm tube coaxial cable, the largest currently used by the Post Office. This has a capacity of 90,000 voice circuits or 18 colour television signals. The optical cable capacity (with eight fibres operating at 140 MBit s^{-1}) is 7,680 two-way circuits or eight colour television signals. It has expansion capacity at higher bit rates to surpass the coaxial cable capacity.

Standard Telephones and Cables

have been obtained in this way that closely follow the form shown in Fig. 4. In principle, this type of measurement could easily be used to study the effects of different launching and fibre lay conditions, so building up families of performance curves. This kind of measurement does not, however, go very far towards explaining why a fibre performs in a particular way. More penetrating measurements are required to achieve such an understanding. In particular, direct measurements of the mechanisms controlling the effects are needed if they are to be thoroughly understood, controlled and exploited.

The simplest study which throws some light on the underlying mechanisms of mode coupling in a multimode fibre is the measurement of the power distribution among the modes as a function of distance, with various excitation distributions. The reduced mode number M equates closely with the angle at which energy leaves the guide when it is terminated with a plane end and allowed to radiate into free space. The semi-angle θ of the cone of radiation from mode number M is given by

$$\theta = \frac{M\lambda}{4a} \quad (5)$$

where a is radius of the core. Thus in a large core fibre, in which the maximum value of M will be large, measurement of the distribution of optical power in the far field gives a simple estimate of the power as a function of mode number. If the fibre is excited by a single mode source coupling predominately to mode $M = 1$, then by cutting back the fibre, the spread of energy out through the mode spectrum can be studied. The measurement is unfortunately destructive but nevertheless does yield valuable information on the type of coupling present and its strength. A variation on this technique monitors the far field distribution while the launching distribution is changed, allowing rather more information to be obtained non-destructively.

A number of other techniques can be brought to bear on the problem, the most powerful potentially being those in which the energy scattered from the fibre is studied as a function of wavelength, angle or distance. Unfortunately such measurements are extremely difficult to make under well enough controlled conditions to obtain useful results. But a related approach promises to allow very sophisticated studies to be made of highly overmoded fibres, namely the recently developed and demonstrated capability to generate and detect very pure single LP modes in fibres capable of supporting tens, hundreds or thousands of such modes. The principle is essentially the same as that involved in launching pure single

modes in planar guides (integrated optics) using a high index prism to couple energy to the fibre by tunnelling from an incident beam through the evanescent field of the cladding into the guided mode core field. Careful control of the angle of incidence and the incident beam profile allows efficient excitation of chosen guided modes of the fibre. This technique was used to obtain the mode shown in Fig. 2.

If the fibre under the prism is maintained parallel, then only modes with significant fields at the cladding–prism interface can be excited efficiently in practice. This means that only those modes close to cutoff can be studied. When the fibre is tapered down under the prism, modes that are well guided in the parallel fibre become accessible at different points along the prism face, so that the whole mode spectrum can be accessed. In either case, the technique can be reversed to allow the power being carried in different modes of the fibre to be separated and studied. For example, in this way the propagation constants of large numbers of discrete, identified modes in a multimode fibre can be measured and information about the index profile deduced. Use of these techniques also opens up the possibility of studying deliberate mode coupling between chosen mode pairs, and some preliminary work of this nature has been reported. The fibre was sandwiched between a pressure pad and a small amplitude grating fabricated on the surface of a sheet of plastic, so that the fibre lay was deformed sinusoidally in a controlled amount at a fixed spatial frequency. The effective frequency could be changed by rotating the grating about its normal axis relative to the fibre. Strong selective coupling was observed between pairs of modes according to the selection rules discussed earlier.

From the point of view of the engineer, enough progress has been made in the fibre and components to make possible serious system design¹. Work is in progress in many laboratories assembling and studying prototype systems, most of which are based on the use of multimode large core diameter fibres. But there remain the finer points to be established, requiring more sophisticated measurement and interpretation of results and only from this second round of studies will the engineering data emerge that are needed for operating systems design. In the meantime the prototype engineering is posing challenges across a broad front to workers of many skills², from materials studies to optical measurement techniques and electromagnetic theory and is forcing a more critical appraisal to be made of previously simple models for fibre propagation.

¹ Marcuse, D., *Theory of Optical Dielectric Waveguides* (Academic, New York, 1974).

² Olshansky, R., and Keck, D. B., *Appl. Optics*, **15**, 48–491 (1976).

³ Gloge, D., and Marcattilli, E. A. J., *Bell Sys. tech. J.*, **52**, 1563 (1973).

⁴ Miller, S. P., Marcattilli, E. A. J., and Li, T., *Proc. IEEE*, **61**, 1703–1751 (1973).

⁵ *IEE Conf. Proc. No. 132, Optical Fibre Communications* (IEE, London, 1975).

Integrated optical circuits for telecommunications

E. A. Ash*, C. W. Pitt* & M. G. F. Wilson*

Optical signals can be guided and manipulated in thin films deposited on a substrate. The new technology will allow the implementation of key circuit functions at optical frequencies.

THERE is no longer any doubt that optical fibre waveguides will shortly be used in telecommunication systems¹. It is an advance which will have an increasing and, in our judgment, dominant role in the development of new communication systems for the remainder of this century. The fundamental advantages of an optical fibre as a communication medium compared for example with a coaxial line must, we feel, eventually prove decisive in many applications. These advantages include the bandwidth, which, at least in principle, is very large; the fact that the cost of the basic fibre material (silica) is negligible as compared with copper—and we are likely to run out of copper before we run out of sand; the almost total lack of interference between different optical fibres, in marked contrast to the situation with coaxial cables, where crosstalk can often represent a daunting problem; the fact that optical fibres are extremely economical of space, occupying an area which is less than one thousandth of that of an equally potent coaxial line.

The technological problems which must be solved before the total potential of optical fibres can be realised are formidable. For reasons which are more fully developed in the preceding article¹, the initial targets will be modest. The transmission medium will consist of fibres with relatively large core diameters (around 50 μm), capable of supporting a large number of optical modes. Since these modes all travel at somewhat different velocities such multimode fibres are rather dispersive, and this limits the bandwidth which can be achieved. Nevertheless it will suffice for the transmission of one or several colour television pictures or the equivalent amount of other forms of data. These early systems will also be restricted to a basic level of topological simplicity (Fig. 1), where a source, be it a laser- or light-emitting diode, is attached to the fibre at one end, and an optical detector at the other. One does not envisage any form of branching circuits. The system is in this sense comparable to the early forms of wired telegraphy over 100 years ago, where each wire had a key at one end and a detector at the other.

When the concept of optical fibres was first seriously discussed, a little over five years ago, it was suggested that perhaps one would never need anything more complex than the linear system of Fig. 1; that fibres would be so small, and so cheap, that one would always prefer to add additional fibres to a system, rather than elaborating the structure in each individual channel. In the intervening years it has become clear that the technological problems of making low loss, reproducible, low dispersion and robust fibres are not yielding without a fight. It is also clear that at least initially it will not prove easy to incorporate a very large number of fibres into a single cable. It is therefore by now accepted that every fibre in a cable will have to be cherished—that it is incumbent on the designers to make optimal use of each. One way in which one could make better use of a fibre than in the system of Fig. 1 is by using the same fibre for both “go” and “return” traffic.

Clearly to do this, one must in some way split the channel into two sections at each end of the fibre. In principle this is readily accomplished using nothing more elaborate than a couple of semi-silvered mirrors as shown in Fig. 2. It is at this point that engineers who have some familiarity with the installation of communication circuits in the field will list some of the painful facts of their lives—such as man-holes, into which the repeaters have to be inserted and which may be half full of water. The use of classical optical components needing careful alignment is not entirely ruled out—but very nearly so.

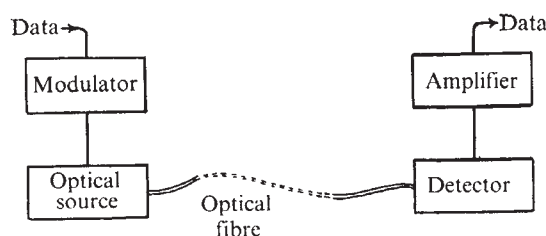
Fortunately we are witnessing the emergence of a new branch of optics—integrated optics²—which shows every sign of being able to perform not only such passive functions as required for the system of Fig. 2, but also a whole range of active functions such as generation, high speed modulation, switching and multiplexing. Integrated optics is lagging behind fibre optics by perhaps as much as a decade. The first generation of optical fibre communication systems will remain within the confines of the system of Fig. 1. But thereafter we see an increasingly important role for integrated optics, in allowing better use to be made of single fibres and in promoting the design of much more complicated and flexible systems. Our purpose in this paper is briefly to review the nature and current status of integrated optics, and then to indicate some of the ways in which it is likely to be incorporated in future optical communication systems.

Thin films

Integrated optics, alias, “thin film guided wave optics”, is concerned with manipulating optical beams guided in a thin film deposited on a substrate having a somewhat lower refractive index. Provided that the film is sufficiently thick it can propagate one or more optical modes. It is also possible to confine such an optical mode laterally, for example by reducing the thickness of the film on either side of the guiding region (Fig. 3). For many purposes it suffices to guide vertically and permit some lateral diffraction of the beam. Typically the film thickness is of the order of 1 μm , a value which allows the use of many of the known techniques for thin film deposition.

Thin films are of course widely used in classical optics for anti-reflection or filter purposes. In these the light propagates normally to the plane of the thin film. In integrated optics propagation proceeds along the thin film. The losses

Fig. 1 Basic optical fibre communication system.



*Department of Electronic and Electrical Engineering, University College London, Torrington Place, London WC1E 7JE, UK.

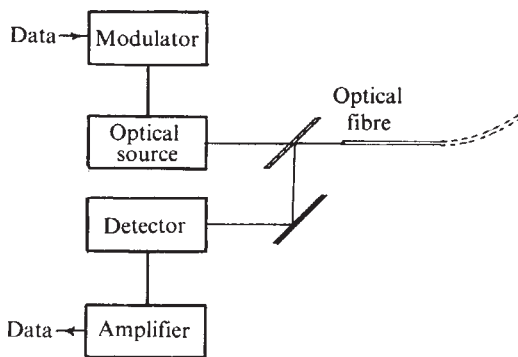


Fig. 2 Terminal for a 2-way optical fibre system, using mirrors for separation of go and return paths.

associated with such films are therefore a much more critical consideration. It has proved possible to achieve attenuation rates of less than 0.1 dB cm^{-1} . In many situations, however, one can work with attenuations of 1 dB cm^{-1} or even higher. These figures are perhaps five orders of magnitude worse than those required, and achieved, in the optical fibre. In integrated optics, however, one is primarily concerned with the loss per wavelength rather than the loss per km, the former figure representing some sort of "Q factor". It turns out that one can make very good components with materials where the fractional loss per wavelength is around 10^{-4} .

Before discussing the materials and other problems encountered in integrated optics, it may be helpful to indicate how one would translate the system of Fig. 2 into thin film form, Fig. 4. The role of the semi-silvered mirror is here played by a Bragg grating structure designed to give a 50:50 split. The laser and detector may be mounted on the optical substrate. It is, however, entirely possible that one would resort to the use of a material based on GaAs, in which case, the laser as well as the detector could be integrated into the structure itself. The overall size of the completed integrated circuit might be 4 mm square. Once made, the components are all rigidly disposed, and no adjustments in the field are required.

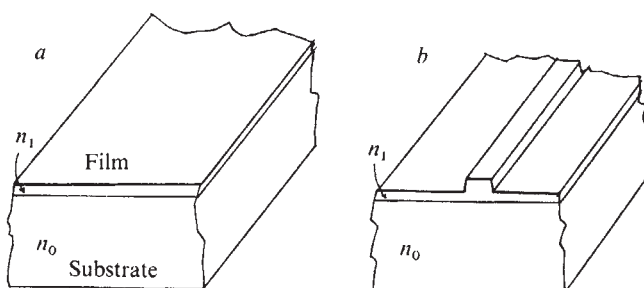
The feasibility of structures of this kind depends in the first instance on the availability of combinations of materials and technologies for depositing the films.

Materials and technology

The selection of materials and deposition technology depends to a large extent on the function of the optical components in the circuit; if devices with laterally defined geometry such as in Fig. 3b are required, the choice of shaping technology will impose further constraints.

The basis of an integrated optical device is light guiding within a thin film. This limits the choice to a film material having a higher index than its supporting substrate. The

Fig. 3 a, Basic integrated optical system. The refractive index n_1 of the film is greater than that of the substrate n_0 . b, Integrated optical system showing one form of waveguide to prevent lateral diffraction effects.



refractive index and thickness of the film will then determine the number and polarisation of modes propagating a transverse plane wave within such a planar guide. In classical optics terms, the refractive index condition ensures total internal reflection of light in the film, while the thickness limitation means that successive reflections of the wave from the thin film surfaces interfere constructively to allow energy to propagate.

A secondary requirement for such guides is an adequately low loss. This may vary from less than 0.1 dB cm^{-1} for an interconnection between components in the circuit, to several dB cm^{-1} for a short active component. Loss, resulting in signal attenuation, derives from three mechanisms: absorption of light in the material, scattering of energy from discontinuities in the refractive index of the film, and scattering from roughness of the film surfaces. The manifestation of these effects varies strongly with the material and the procedure adopted for film formation. For example, a dopant such as copper can be introduced by diffusion into a single crystal material such as magnesium oxide to form a waveguide; its effect will be to increase the real part of the refractive index but it will have little effect on the absorption (the imaginary part of the index). If, however, a few parts per million of copper are inadvertently introduced into a sputtered glass waveguide

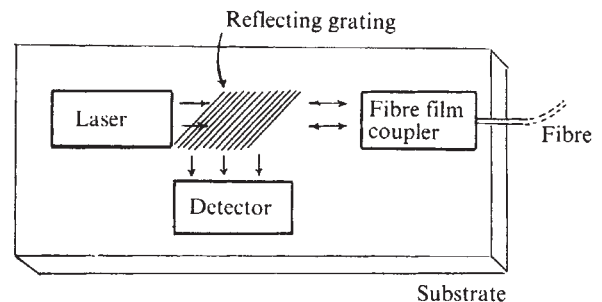


Fig. 4 Integrated form of two-way communication terminal.

it results in severe absorption. The effect of bulk and surface perturbations of the index and topography of the guide depend on the magnitude of these effects in relation to the wavelength. Scattering of an electromagnetic wave from a randomly rough surface is negligible if glancing incidence occurs, provided that the wavelength is large compared with the roughness amplitude, or that the correlation length of the perturbation is greater than a few tens of wavelengths.

These considerations of the material specifications bear on the basic problem of guiding; but we also need to manipulate the optical signal. Whereas some functions may be achieved with passive components, such as lenses, diverging guides and gratings, the realisation of a light source, modulation, switching or detection function requires the use of active materials. These may be lasing, electro-optic, piezoelectric, or photoelectric sections of the circuit. Such an additional degree of freedom for the circuit designer must, however, be paid for in an increased fabrication complexity. For example it may be desirable to phase modulate a signal in a zinc oxide film on a glass substrate. The electro-optic coefficients of ZnO are relatively large, producing a suitable change of refractive index when an electric field is applied across the guide. The appropriate crystal orientation (c axis parallel to the applied field) may be achieved by sputtering the material in appropriate conditions³: the film structure is polycrystalline. At typical deposition rates for such a film, the crystalline column dimensions are unfortunately similar to the wavelength of visible light, and prove to be efficient scatterers so that

Table 1 Materials and technology of integrated optics

Substrate material—refractive index and crystalline state	Film Material—refractive index and crystalline state	Operational wavelength range	Film deposition technique	Film shaping process	Guide attenuation dB cm^{-1}	Component class
Glass Soda-lime silicate glass $n_{0.63\mu\text{m}} = 1.5124$ Amorphous	Glass Aluminium borosilicate $n_{0.63\mu\text{m}} \approx 1.52$ –1.58 Amorphous	Visible to near infrared	Sputtering	Ion etching	0.5–2.0	Passive
	Organics Photoresists, fatty acids, polymers. $n_{0.63\mu\text{m}} \approx 1.5$ –1.6 Amorphous and planar ordered	Long ultraviolet to visible	Solvent deposition, Langmuir process, Chemical vapour phase.	Selective polymerisation, or scission (direct lithography by irradiation)	0.1–2.0	Passive
	Metal oxides for example Al_2O_3 , Ta_2O_5 , Nb_2O_5 $n_{0.63\mu\text{m}} \approx 1.60$ –2.24 Amorphous and polycrystalline	Visible to near infrared	Electron-beam evaporation, Sputtering, Chemical vapour phase	Ion or chemical etching	~0.5–5.0	Passive and active
	Doped glass Ag substituted sodium glass $n_{0.63\mu\text{m}} \approx 1.6$ Amorphous	Ultraviolet to visible	Ion exchange	Masking of non-guiding region	~1.0	Passive
Crystal dielectric Lithium niobate $n_{0.63\mu\text{m}} = 2.20$ Single crystal	Doped crystal Titanium dopant $n_{0.63\mu\text{m}} = 2.22$ –2.25 Single crystal	Visible and near infrared	Evaporation (of dopant) and diffusion	Confinement of diffused metal to guiding region	0.1–1.0	Active
	Titanium dopant $n = 2.25$ Partially disordered crystal	Visible and near infrared	Ion implantation	Confined ion-beam or masking	0.5–5.0?	Active
	Lithium depleted $n = 2.21$ –2.22 Single crystal	Visible and near infrared	Out diffusion of lithium	Passive masking (to prevent egress of lithium)	1.0	Active
Semiconductor Gallium arsenide $n_{0.905\mu\text{m}} = 3.34$ Single crystal	Graded doped crystal GaAsP $n_{0.9\mu\text{m}} = 3.4$ Single crystal	Near infrared	Vapour phase epitaxy	Chemical or ion-etching	1–5	Active
	Proton irradiated GaAs $n = ?$ Disordered crystal	Near infrared	Ion implantation	Chemical or ion-etching	~8	Active
	Epitaxial multilayer GaAlAs/GaAs $n_{1.0\mu\text{m}} \approx 3.3$ Single crystal	Near infrared	Molecular beam epitaxy	—	~10	Active

the attenuation approaches 100 dB cm^{-1} . By careful modification of the process, the crystalline dimension may be reduced to rather less than the wavelength, with a corresponding reduction of the loss to about 8 dB cm^{-1} .

It is inevitable that the integrated optical circuit will introduce significant additional attenuation which it is important to minimise. One method of achieving this aim is to reduce the length of interconnection links in the circuit. This implies a capability to direct the light around small radius bends. Stripe waveguides meet these needs rather better than planar guides, but at the expense of an additional shaping operation, such as ion machining the planar films or restricting the film deposition to cover a predetermined path. Additional limitations are also imposed on material selection.

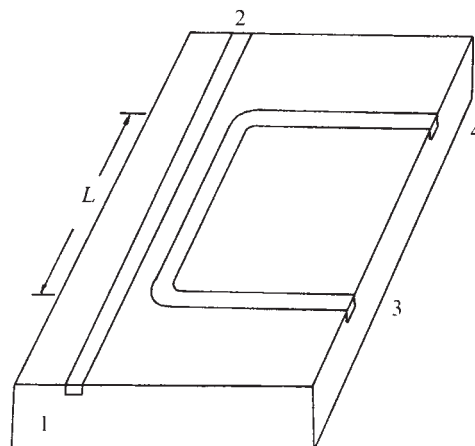
A prime requirement of a stripe guide is tight confinement of the energy within the stripe; unfortunately strong confinement implies a large index difference between the guide and surrounding media and this is the worst configuration for loss of energy from guide surface roughness. The introduction of two additional surfaces to the guide further degrades the attenuation due to scattering.

In the table we illustrate some typical material combinations with appropriate technologies for deposition and shaping. The technological selection problems may be illustrated by examining two of these systems which are currently used for fabrication of experimental devices.

The titanium doped lithium niobate⁴ waveguide has found wide application because it combines a low optical loss with strong electro-optic and piezoelectric properties. The low absorption loss of the titanium ions may be attributed to

their uniform dispersal throughout an almost defect free crystal lattice and the fact that they are in an oxidation state in which they exhibit small electron-photon interaction. Further, the waveguide is diffused resulting in reduction of scattering loss from the substrate-film interface (in contrast to a discrete film guide)—and also in the reduced edge acuity problems associated with defining the lateral boundaries of a stripe waveguide. The resulting waveguide has, however, a high refractive index and this leads

Fig. 5 Directional coupler; power inserted at arm 1 emerges from arms 2 and 4. The electro-optically controlled coupler incorporates electrodes on either side of guide 1–2 over the length L .



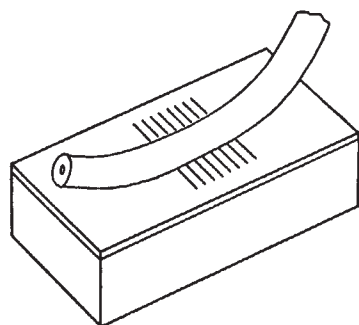


Fig. 6 Fibre-film coupler showing grating necessary for phase matching.

to a large scattering loss at the air-guide interface—which fortunately may be reduced by polishing. Less tractable is the problem of coupling guided energy inadvertently to the high refractive index substrate when a low refractive index medium, such as an optical fibre, is brought into contact to transfer energy to the fibre.

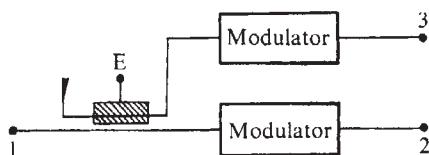
The GaAs material systems^{5,6} are obviously attractive as they enable a lasing source to be interconnected through GaAs/GaAlAs waveguides to other active components made from the same materials and integrated on to one substrate. There are, however, two major problems. GaAs lasers operate at a wavelength close to the fundamental absorption edge of the material and proton-implanted or epitaxial GaAs guides exhibit a 'smeared' absorption edge resulting from damage or dopants within the guiding region. The broadening of the absorption edge extends into the lasing wavelength region resulting in the rather high absorption losses shown in the table. The other problem is similar to that encountered in lithium niobate—how to couple most of the laser energy from the high index GaAs circuit into low index fibre without incurring unwanted substrate interactions.

Functional elements

Although the name integrated optics accurately reflects the concept, little work has yet been done on integrating many components on to a single substrate. Research has concentrated on individual circuit elements, often derived directly from their counterparts in microwave engineering, whereas others resemble more closely components in bulk optics. A complete integrated optics circuit may include waveguides, power dividers, amplitude and phase modulators, switches, deflectors, filters, sources and detectors. To illustrate the current problems and state of the art, we shall briefly discuss three key components: directional couplers, fibre-film couplers and lasers.

A key component in any guided wave system is the directional coupler. It enables energy to be coupled, selectively, from one waveguide to another; its counterpart in conventional, free space optics is the partially reflecting mirror. The basis for directional coupling is the interaction of two waveguide modes having equal, or very nearly equal velocities. A particular configuration is shown in Fig. 5 where two waveguides are brought into close proximity. The fields of one guide overlap the adjacent guide; even though this overlap may be small, large transfers of power can occur over the interaction length L by virtue of

Fig. 7 Multiplexing circuit. Division of power from port 1 between ports 2 and 3 is controlled by the switchable directional coupler, terminal E.



the synchronism in velocities. Interaction lengths are of the order of 1,000 wavelengths so that a slight deviation from velocity synchronism will result in a large cumulative phase error between the coupled modes, leading to a serious degradation in performance. It follows that the tolerance restrictions on dimensions and material parameters are severe. Dimensional tolerances of less than 0.1% are typical. In spite of this, successful couplers have been fabricated using electron beam lithography. Furthermore, the tolerances can be relaxed considerably if the velocities are suitably tapered along the direction of propagation.

Both the routing and modulation of optical signals can be performed with the help of electrically controlled switches. Several types of switch have been described which rely for their action on electroabsorption, Bragg diffraction and other well known processes; but in spite of different geometries most of them rely on a change of refractive index induced by electro-optic or acousto-optic effects. As an example, consider once again the directional coupler shown in Fig. 5. Suppose it to be fabricated in an electro-optic material with electrodes arranged so that an electric field can be applied across one of the coupled guides. With no applied field the coupler is designed so that all the power entering arm 1 is completely coupled into the adjacent guide and emerges from arm 4. When an electric field is applied to the electrodes the refractive index in one guide is changed, the mode synchronism destroyed and the power transfer reduced to a very small fraction, so that almost all of the input power emerges from arm 2. Although the obtainable changes in refractive index are small, of the order of 10^{-4} , the tolerance on synchronism is so severe that only

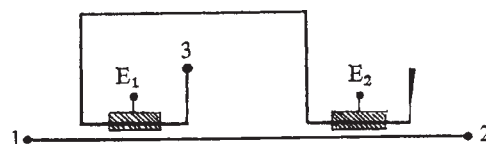


Fig. 8 Optical circuit for interconnecting any two of the three ports, 1,2,3 by means of two directional couplers, E_1 , E_2 .

a small change in index is sufficient to 'spoil' the interaction.

Integrated optics will only play its full part in optical fibre communication systems when efficient fibre-film couplers become available. There are two main approaches to solving this problem. In the first the fibre is butt jointed to the end of the film. This method is the most direct and simplest in concept, but coupling is limited to the periphery of the circuit and the end faces of both film and fibre must be of high optical quality. The fibre must be aligned and held in position with respect to the film within a fraction of the core diameter. For efficient coupling the shape of the film and fibre fields should be matched as closely as possible.

To avoid some of these difficulties an alternative approach has been attempted: the fibre is pressed on to the film at any point on the circuit as shown in Fig. 6. Distributed coupling is achieved as in the directional couplers, the precise form of the fields is unimportant so long as overlap occurs. The grating shown in Fig. 6 is necessary to accommodate the possible large difference in velocity between fibre and film modes. The device is a directional coupler and the power transfer is determined by the interaction length and strength of coupling.

There is a wave velocity tolerance problem but again this can be overcome by suitable tapering. The fibre cladding must be partially removed to expose the core field to the film but this can be done by careful experimental techniques. A much more serious problem arises from the difference in index between film and fibre. This is especially marked when the substrate index is higher than the fibre



Switchboard of the Royal Exchange Manchester, 1895.

Post office

index; it is this situation which is encountered when coupling a glass fibre to a doped lithium niobate film. It leads to a large fraction of the light being coupled to the substrate rather than to the film. Recent attempts to overcome this have relied on the deposition of even higher index materials on to the waveguide film but this brings us back to one of the original problems in integrated optics—the deposition of low loss high refractive index films. Thus the design of an efficient coupler is intimately bound to the materials problem.

Although most research work on semiconductor injection lasers proceeds independently of the work on integrated optics, there is a significant effort to produce thin film lasers which can be fabricated within an integrated optics thin film circuit. Consider the particular problem of fabricating thin film end reflectors. In the conventional injection laser these are provided by cleaving the end faces of the crystal. The integrated optics approach is to etch gratings having a pitch equal to one half wavelength of light so that the Bragg condition results in light being reflected normally from the grating array. The device thus comprises two sets of gratings on either side of the lasing film. In the distributed feedback laser the lasing action occurs within a single corrugated film obviating the need for two separate gratings.

It is interesting to note that integrated optics techniques can be used within the cavity of a conventional injection laser. This approach allows the use of a separate intracavity modulator which has significant advantages over direct laser modulation.

Integrated optics for fibre systems

We have already seen one example in Fig. 4 of an application of an integrated optical system permitting two-way communication along a single optical fibre. In this we assume that the transmitting laser at each end would be of the same frequency. The required 50:50 Bragg reflector in Fig. 4 implies a power loss of one half at each end, that is 6 dB, or three-quarters overall. It is a modest but not insignificant price to pay for being able to double the usage of the fibre. Since the operating wavelength of injection laser can be tightly controlled, one could also envisage a system in which the two lasers are separated by a small frequency interval. If the two laser frequencies differ, even by as little as 1%, it is possible to design the Bragg reflector so as to

make it selective. It is then no longer necessary to incur this 6 dB loss.

Optical fibres have an optimum wavelength of operation, but in most cases this optimum is rather flat. Now the spectral width of injection lasers is very large as compared with other solid state or gas lasers, but is still typically no more than 0.2%. If the fibre performs adequately over a 20% frequency band, this would imply that one could have up to 100 separate frequencies propagating along the single fibre. The required channel dropping filters at each end of the fibre could be implemented by a series of tuned Bragg reflectors or using holographic couplers. Even if one were to implement this idea in a relatively modest way—perhaps 4 channels each way—it might well prove to be a better way of increasing the data capacity of the fibre than attempts to use very short pulses, which, quite apart from the dispersion problems, imply the need for very high speed electronics.

The inverse problem will also arise where the need is to feed several fibres from a single source as shown in Fig. 7. Here the power from the laser can be directed to either of the two fibres, or indeed split in some required proportion. Each channel will then be separately modulated by the electro-optic modulator. The same basic circuit can of course also be used to multiplex a single detector.

One can also anticipate situations where one would wish to interconnect three fibres, in such a way that any two can be interconnected. Figure 8, shows how this can be achieved using only two switchable directional couplers. The basic unit could of course be combined with others to realise more complex switching systems.

A final example is provided by the requirement for a data highway, where information can be injected on to or extracted from a fibre which may run throughout a large system—such as an aircraft or ship. Again the switchable directional coupler provides a solution.

The integrated optical solutions to these systems problems display a measure of elegance; the basic technology for their implementation exists. This might then raise the question as to why we anticipate that there will be a substantial delay before they are used in practice. Of course there is always a large gap between a technology which has passed a feasibility test, and one on which one can place the reliance which is essential in a telecommunications context. There is, however, an additional reason. As indicated by Midwinter¹, the first fibre systems will certainly use multimode fibre. Almost everything that we have said about integrated optical components and their application has tacitly assumed that we are concerned with a single mode optical fibre system. There is no doubt that single mode systems will arrive on the scene eventually—the technology is harder, but in every other way they are simply much better. We therefore foresee that the main applications for integrated optics will not make an impact until that time.

It would, however, be wrong to give the impression that integrated optics has no role in multi-mode systems. There are functions which can be performed even if one is faced simultaneously with a large number of modes. For example, we are currently developing switches and tapping points to be used in conjunction with multimode fibres. The attainable performance will fall short of that which can be anticipated for single mode systems but we nevertheless expect the devices to be of use. For the dawn of the golden age of integrated optics, however, we must await the arrival of single mode optical fibre systems.

¹ Midwinter, J. E., *Nature*, **261**, 371–376 (1976).

² Tamir, T. (ed.), *Integrated Optics, Topics in Applied Physics* (Springer-Verlag, 1975).

³ Pitt, C. W., Gfeller, F., and Stevens, R., *Thin Solid Films*, **26**, 25–51 (1975).

⁴ Schmidt, R. V., and Kaminow, I. P., *Appl. Phys. Lett.*, **25**, 458–460 (1974).

⁵ Sopori, B. L., and Chang, W. S. C., *Topical Meeting on Integrated Optics, 1976* (Optical Society of America, Washington, DC).

⁶ Merz, J. L., and Cho, A. Y., *Topical Meeting on Integrated Optics, 1976* (Optical Society of America, Washington, DC).

articles

r-Process nucleosynthesis of superheavy nuclei and nuclear mass tables

G. J. Mathews & V. E. Viola, Jr

Department of Chemistry, University of Maryland, College Park, Maryland 20742

New predictions of nuclear masses and fission barriers indicate that it is not possible to form superheavy nuclei during r-process nucleosynthesis.

THE r process is the sequence of rapid neutron captures and subsequent β decays responsible for the build up of many of the heavier nuclei in nature¹. In particular, the r process is probably the only natural mechanism for the formation of nuclei with $Z \geq 83$ and $A \geq 210$. Because of this, there has been considerable speculation concerning the possibility of forming superheavy nuclei ($A \approx 290$ –300) in the r-process environment^{2–10}. Here we summarise the results of an intercomparison of the r-process paths predicted by several of the more recent formulations for nuclear masses far from the line of β stability. These predictions include detailed empirical systematics based on nuclear decay energies¹¹, and the addition of higher order terms to the liquid-drop mass formula important for r-process nuclei (ref. 12, and W. D. Myers, unpublished). In addition to being the most up-to-date masses available, these masses have been derived by independent methods; on the one hand we have strictly empirical extrapolations, and on the other hand, predictions based on models with a minimum number of free parameters. Because these methods give the two extremes of nuclear stability, we believe that these mass tables should bracket the actual r-process nuclear masses. We find that the newer nuclear masses consistently preclude the possibility of forming superheavy nuclei during most of the r process.

r Process and nuclear masses

Nuclear masses enter into a calculation of the r process in two important ways. First of all, the neutron binding energies are used in the determination of the r-process path. In a static r-process calculation² one assumes an $(n, \gamma) \rightleftharpoons (\gamma, n)$ equilibrium. The high neutron flux responsible for this equilibrium presumably arises from violent stellar processes such as a supernova. The Saha equation¹³ can then be applied to describe the statistical equilibrium between neutron emission and capture. Thus, the relative isotopic abundances are given by

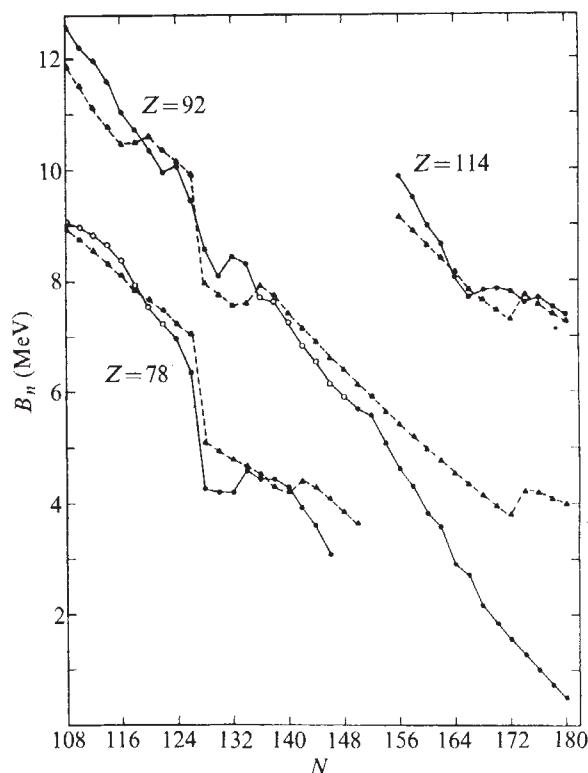
$$\frac{\eta_n n(Z, A-1)}{n(Z, A)} = \frac{(2\pi\mu k T)^{3/2}}{h^3} \left[\frac{G_n G(Z, A-1)}{G(Z, A)} \right] \exp \left[\frac{-B_n(Z, A)}{kT} \right]$$

where η_n is the neutron number density; $n(Z, A)$ is the relative abundance of isotope A with atomic number Z ; $G(Z, A)$ is the partition function for the nucleus A_Z ; T is the r-process temperature, and $B_n(Z, A)$ is the neutron binding energy for the nucleus A_Z . The exponential makes the r-process abundances very sensitive to neutron-binding energies. As a consequence, one obtains a maximum r-process abundance path characteristic of each set of nuclear masses (see Figs 3–5). This path is centred

along some optimum value of the neutron binding energy (typically ~ 3 MeV), and relative abundances become insignificant within several nucleon numbers of this optimum.

A second important influence of nuclear masses on an r-process calculation arises when one attempts to determine the values of Z and A where neutron induced fission begins to dominate neutron capture. This is the point at which the build up of heavier elements in the r process ceases. It has been shown⁸ that, to a good approximation for r-process purposes, all nuclei can be said to undergo neutron induced fission whenever the sum of the incident neutron kinetic energy and the neutron binding energy exceed the height of the fission barrier. The determination of the (n, f) cutoff is then reduced

Fig. 1 Comparison between empirical and mass formula neutron binding energies for $Z = 78, 92$ and 114 . It is seen that the empirical estimates tend to predict less stability for neutron-excess isotopes. The neutron binding energies for the Seeger and Howard masses (not shown) exhibit behaviour similar to the Myers and Swiatecki masses. $\circ$, Measured values²⁴; $\bullet$, ref. 11; $\blacktriangle$, Myers (unpublished).



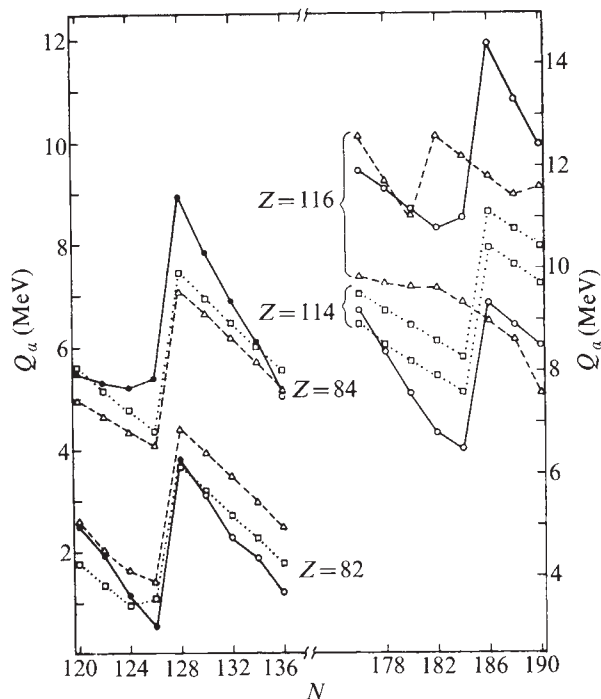


Fig. 2 Comparison between empirical and mass formula α -decay energies. It is seen that the effects of shell closures are more dramatic for the empirical masses. . . ., Myers (unpublished); ----, ref. 12; —, ref. 11.

to the estimation of fission barriers and neutron binding energies, both of which are sensitive to the mass formula used.

Nuclear masses

We have used three different compilations of nuclear masses. Empirical estimates of the binding energies of r -process nuclei have been derived by Viola *et al.*¹¹ on the basis of α - and β -decay energy systematics. Since these are rather smooth except near closed shells and since the experimentally known α -decay chains span a rather broad range of nuclei, one can make extrapolations into the r -process region with some degree of confidence. Predictions based on α -decay systematics have been shown, for example, to yield rather impressive agreement with experimental α -decay Q -values of neutron-deficient nuclei¹⁴.

The Seeger and Howard mass formula¹² involves the usual separation into a smooth macroscopic liquid drop term and microscopic corrections. The expression for the macroscopic energy includes shape-dependent surface and surface asymmetry terms, as well as a shape-dependent Coulomb term corrected for a diffuse nuclear surface. The microscopic corrections employ the Strutinsky renormalisation procedure and BCS pairing theory. The parameters of the formula are adjusted to give, simultaneously, self-consistent nuclear masses and fission barriers. The fission barriers are calculated assuming the idealised liquid drop saddle point shapes of Cohen and Swiatecki¹⁵.

The third formulation of nuclear masses is the droplet model formula of Myers and Swiatecki¹⁶. In addition to containing many of the features of the liquid drop approach, the droplet model allows the neutron and proton densities to vary independently, and includes higher order terms in $(N - Z)$ which are important for the neutron-excess r -process nuclei. The new droplet model parameters, as in ref. 12, are determined from a fit to experimental masses and fission barriers.

In general the masses derived from empirical systematics predict increased stability for neutron-deficient isotopes and decreased stability for neutron-excess isotopes compared with the predictions of the mass formulae. This is demonstrated in

Fig. 1 where we compare trends in neutron binding energies for $Z = 78, 92$ and 114 . The binding energies for both mass formulae show similar behaviour. As far as the r -process is concerned, this means that the r -process path for the masses based on empirical systematics will follow a less neutron-rich path than those of the mass formulae.

Also, the masses based on empirical systematics predict much stronger discontinuities in nuclear structure in the region of the closed shells. For example, in Fig. 2 we compare the α -decay Q -values for even-even nuclei near the $Z = 82, N = 126$ doubly closed shell. It is observed that the mass formulae predict much weaker shell strengths than are actually observed. This tendency for the mass formulae to underestimate shell corrections is an important consideration in the decay-back process after r -process freeze out. It implies that real nuclei just beyond the $N = 184$ closed shell should be more unstable than the mass formulae predictions for both α decay and, more particularly, spontaneous fission. Since many of the progenitors of the superheavy nuclei that one might expect to observe in nature must β decay through $N > 184$ species to reach the 'island of stability', this factor will further inhibit survival in the decay-back process.

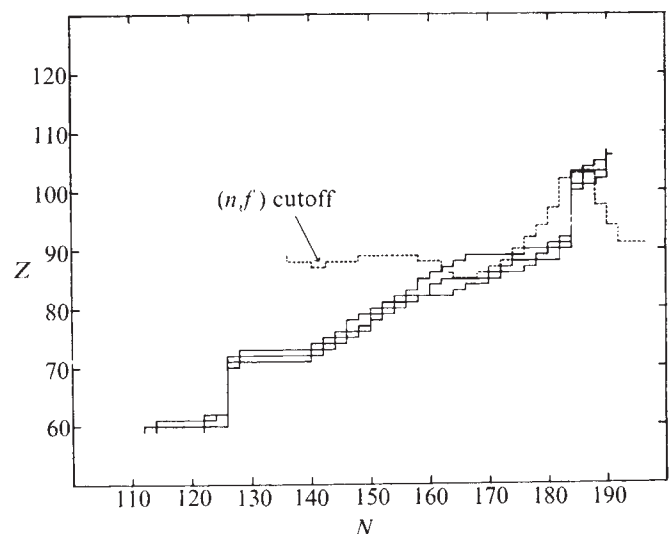
Results of calculation

To compare the different sets of nuclear masses we have performed a static r -process calculation, as outlined in ref. 2, for each mass table separately. The r -process temperature and density (or equivalently the r -process cycle time²) were then adjusted to reproduce the centroid of the experimentally observed $A = 195$ abundance peak. β -decay rates were estimated using the formulation of Senbetu¹⁷ for calculation of f and assuming⁶ an average $\log ft$ value of 6.5.

Figures 3–5 show the r -process paths, conditions, and the neutron-induced cutoff line estimated on the previously mentioned procedures. The path which best fits the $A = 195$ abundance peak is given by the central solid line on each figure. The paths on either side of the 'best fit' line correspond to a displacement of the centroid of the $A = 195$ peak by one mass unit. This was taken as the limit of uncertainty in the calculation.

Figure 3 shows the r -process path predicted for the Myers droplet model masses. We find that the 'best fit' r -process path crosses the (n, f) cutoff line at $Z = 85$ and $N = 164$. Although it appears at first glance that one might, with a slightly more neutron-excess path, temporarily escape the neutron-induced fission cutoff until $Z = 102$, this is actually not the case. As has been pointed out by Schramm and Fiset⁸, relative abundances of a fission-unstable isotope as low as $\sim 10^{-10}$ are sufficient to

Fig. 3 r -Process path and fission cutoff for the Myers (unpublished) mass formula, using $T = 1.8 \times 10^9$ K, and $\eta_n = 1.5 \times 10^{26}, 6.0 \times 10^{26}$ and 2.2×10^{27} n cm⁻³.



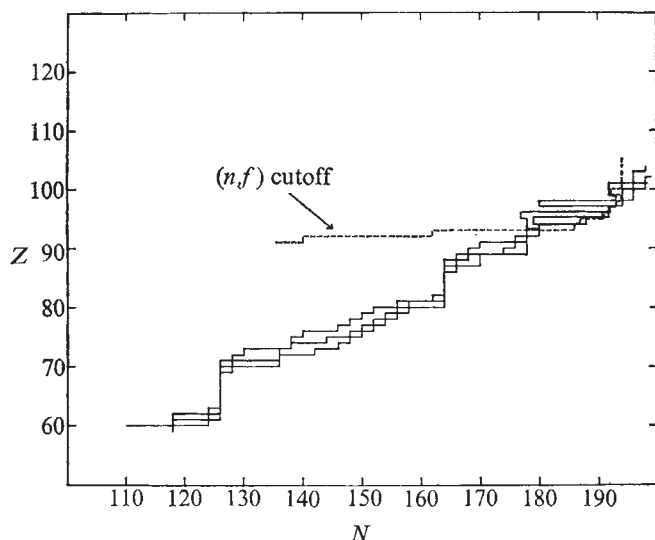


Fig. 4 r-Process path and fission cutoff for the Seeger and Howard mass formula¹², using $T = 1.8 \times 10^9$ K and $\eta_n = 1.9 \times 10^{27}$, 7.5×10^{27} and 3.4×10^{28} n cm⁻³.

terminate the r process. This is so because the neutron capture rate, and thus the (n,f) rate, is so much more rapid than the β -decay rate: for example $\lambda_{\beta}/\lambda_n \approx 10^{-10}$ for $T \approx 10^9$ K, $\eta_n \approx 10^{28}$ n cm⁻³. To reduce the abundances of isotopes along the (n,f) cutoff line sufficiently, an extremely neutron-excess path is required, which we feel is outside the limits of uncertainty for the present r-process calculation.

Figure 4 shows the r-process path predicted for the masses of ref. 12. Here we find that the (n,f) line is crossed by the best fit r-process path at $Z = 93$ and $N = 178$, again short of the superheavy regime.

In Fig. 5 the r-process path predicted by the empirical masses of Viola *et al.*¹¹ is plotted. For our calculations using these masses, we took the fission barriers from ref. 12 since these give the best fit to experimental barriers. The empirical α -decay systematics do not extend below $Z = 50$, and the r-process calculations generally extend to $Z \approx 40$; therefore, the Seeger and Howard masses were used to fill in the r-process calculation for the empirical masses. The results were not particularly sensitive to the choice of masses below $Z = 50$, since all of the mass formulae seem to predict similar behaviour in this region.

Fig. 5 r-Process path for empirical masses of Viola *et al.*¹¹. The fission barriers for the (n,f) cutoff line are assumed from the tables of Seeger and Howard¹². $T = 1.8 \times 10^9$ K; $\eta_n = 3.3 \times 10^{27}$, 1.9×10^{28} and 3.3×10^{29} n cm⁻³.

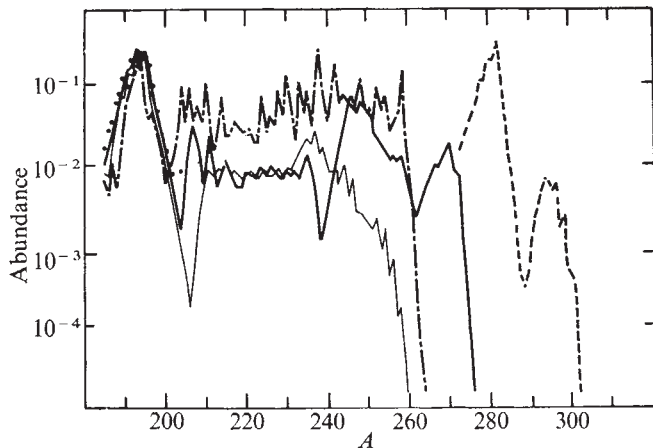
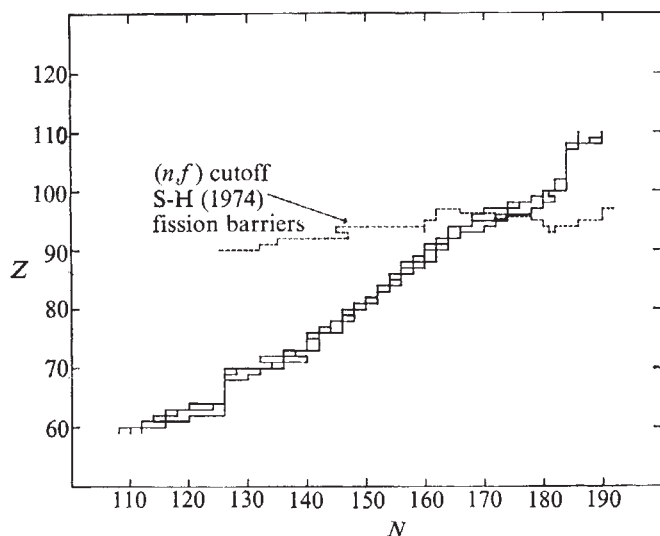


Fig. 6 r-Process abundances before decay-back predicted for the best fit to the $A = 195$ abundance peak for each mass formula., r-Process abundances; —, Myers (unpublished); — · —, ref. 12; - - - -, ref. 11; · · · · ·, ref. 16.

The abundances before decay-back of the r-process paths which best fit the $A = 195$ peak for each mass table are shown in Fig. 6. Also shown for comparison are the abundances calculated with the Myers and Swiatecki liquid-drop formula¹⁸. For this mass formula, a maximum Z of 100 before fission termination was assumed from the work of Schramm and Fiset⁸. Although previous work seemed to indicate that the r process could form the minimum mass number necessary to reach the island of superheavy nuclei⁶, the newer atomic mass compilations provide a much more pessimistic conclusion.

The primary reason for the difficulty in reaching the superheavy nuclei in our calculation is the increased magnitude of the surface asymmetry term in the newer mass formulae. The importance of this term in r-process calculations has been pointed out by a number of authors^{5,8,10}. This increase in the magnitude of the surface asymmetry term results in lower fission barriers along the r-process path and, hence, neutron-induced fission becomes important at lower values of Z and A . A secondary reason for the decrease in the maximum mass number achievable is that the newer mass tables predict slightly less neutron-excess r-process paths near the fission cutoff. This effect is most significant for the empirical masses, and is consistent with the predictions based on the energy-density mass formula of Brueckner *et al.*⁹. A consequence of these less neutron-excess r-process paths is that even if fission barriers were not reduced to the extent implied by the newer mass formulae, one would still encounter difficulty in obtaining mass numbers of $A \gtrsim 290$ needed to form the superheavy nuclei.

The fission cutoffs shown in Figs 3–5 might be considered as upper limits rather than actual limits for the following reasons:

- (1) We have not considered here the problem of zero point motion which could tend to lower the effective fission barriers as pointed out by Schramm and Fiset⁸.
- (2) We have treated the neutron kinetic energies as a delta function at 100 keV. In reality, as pointed out by Johns and Reeves¹⁸, one would expect a Maxwell-Boltzmann kinetic energy distribution for the neutrons. This would result in a large number of higher energy neutrons which could induce fission earlier in the r process.
- (3) In these calculations we have not introduced a correction to the Coulomb energy from the interaction between the nucleus and the electronic charge in a dense stellar plasma. It has been estimated that this interaction should significantly reduce the effective fission barriers for r-process nuclei¹⁹.

Conclusion

It seems that the most recent nuclear mass compilations make the possibility of the formation of superheavy nuclei during

most of the r-process quite difficult. Although in a dynamic calculation one might expect to form a small amount of super-heavy nuclei during 'freeze out' (ref. 20 and O. Johns, unpublished) the bulk of the r-process period (responsible for the $A = 195$ abundance peak) seems not to produce superheavy nuclei, according to the most recent predictions of nuclear stability. This is then, perhaps, one reason why experimental searches have yielded rather stringent upper limits on the abundances of superheavy nuclei in nature²¹⁻²³.

We thank P. A. Seeger and W. D. Myers for furnishing the mass tables used. We also thank the US Energy and Research Development Administration and the University of Maryland Computer Science Center for support.

Received January 15; accepted March 22, 1976.

- ¹ Burbidge, E. M., Burbidge, G. R., Fowler, W. A., and Hoyle, F., *Rev. mod. Phys.*, **29**, 547 (1957).
- ² Seeger, P. A., Fowler, W. A., and Clayton, D. D., *Astrophys. J. Suppl. No.* 97, **11**, 121 (1965).

- ³ Viola, V. E., *Nucl. Phys.*, **A139**, 188 (1969).
- ⁴ Berlovich, E. E., and Novikov, Yu. N., *JETP Lett.*, **9**, 445 (1969).
- ⁵ Schramm, D. N., and Fowler, W. A., *Nature*, **231**, 103 (1971).
- ⁶ Fiset, E. O., and Nix, J. R., *Nucl. Phys.*, **A193**, 647 (1972).
- ⁷ Onishi, T., *Prog. theor. Phys. Osaka*, **47**, 845 (1972).
- ⁸ Schramm, D. N., and Fiset, E. O., *Astrophys. J.*, **180**, 551 (1973).
- ⁹ Brueckner, K. A., et al., *Phys. Rev. C*, **7**, 2123 (1973).
- ¹⁰ Howard, W. M., and Nix, J. R., *Nature*, **247**, 17 (1974).
- ¹¹ Viola, V. E., Swant, J. A., and Graber, J., *At. Data Nucl. Data Tables*, **13**, 35 (1974).
- ¹² Seeger, P. A., and Howard, W. M., *Nucl. Phys.*, **A238**, 491 (1975).
- ¹³ Clayton, D. D., in *Principles of Stellar Evolution and Nucleosynthesis*, ch. 7 (McGraw-Hill, New York, 1968).
- ¹⁴ Viola, V. E., in *Proceedings of the Third International Conference on Atomic Masses* (edit. by Barber, R. C.), 135 (University of Manitoba Press, Winnipeg, 1967).
- ¹⁵ Cohen, S., and Swiatecki, W. J., *Ann. Phys.*, **19**, 67 (1962).
- ¹⁶ Myers, W. D., and Swiatecki, W. J., *Ann. Phys.*, **55**, 395 (1969); **84**, 786 (1974).
- ¹⁷ Senbetu, L., *Phys. Rev. C*, **7**, 1254 (1973).
- ¹⁸ Johns, O., and Reeves, H., *Astrophys. J.*, **186**, 233 (1973).
- ¹⁹ Duorah, K., and Duorah, H. L., *Phys. Rev. C*, **8**, 725 (1973).
- ²⁰ Schramm, D. N., *Astrophys. J.*, **185**, 293 (1973).
- ²¹ Behringer, K., Grütter, A., von Gunten, H. R., Schmidt, A., and Wyttenbach, A., *Phys. Rev. C*, **9**, 48 (1974).
- ²² Popeko, A. G., Skobelev, N. K., and Ter-Akopyan, G. M., *Phys. Lett.*, **52B**, 417 (1974).
- ²³ Anders, E., and Larimer, J. W., *Science*, **175**, 981 (1972).
- ²⁴ Wapstra, A. H. and Grove, N. B., *Nuclear Data Tables*, **9**, 265 (1971).

Environmental and climatic implications of late Quaternary lake-level fluctuations in Africa

F. Alayne Street & A. T. Grove

Department of Geography, Downing Place, Cambridge CB2 3EN, UK

Widespread water balance shifts in closed basin lakes reveal major glacial-interglacial variations in rainfall and streamflow. Markedly arid conditions characterising inter-tropical Africa during the last glacial maximum were followed by an early Holocene lacustral phase in which annual precipitation between 24°N and 8°S exceeded 150% of the 1931-60 mean.

We have compiled published lake-level data from 58 African basins in a simple cartographic form which emphasises regional patterns of surface water availability over the past 21,000 yr. With certain reservations, these maps can be used to interpret past precipitation-evaporation patterns. We hope that our approach will stimulate discussion on the long term shifts in atmospheric wind systems and moisture transport over Africa and surrounding oceans. The results also relate directly to past erosion rates and geochemical cycling, to the timing of aquifer recharge, to the distribution and spread of flora and fauna, and to the diffusion of prehistoric cultures over the continent. We suggest that they will explain many anomalies in the patterns of water-borne diseases. Since significant parallels exist with the lake-level chronologies of other regions, we hope to collaborate with other workers in extending this series of maps to other continents.

Africa has two advantages for this initial survey: a large land area spanning over 60° of latitude, where direct hydrological effects of glaciation were minimal; and the best coverage of radiometrically-dated studies of lake levels¹⁻⁷. It has long been the scene of a lively debate on the question of the synchronism of glacial and 'pluvial' periods⁸. Changes in the water level have been inferred not only from former strandlines and overflow channels, but also from the depositional history of cores, exposed sections and archaeological excavations.

Technique

Relying exclusively on radiocarbon-dated chronologies, we have constructed a series of maps representing lake-level

patterns during each millennium since 21,000 yr b.p. Selected examples appear in Figs 1-5. We have specifically excluded all relative dates, all archaeological information unrelated to shorelines, and inferences about past climates derived from regional vegetation or slope processes. Evidence from alluvial terrace sequences has not been plotted. We have also avoided lakes which developed by blockage of major throughflowing rivers like the Niger and the Nile.

In Figures 1-5 we use the tripartite division of lake-level curves devised by Butzer *et al.*². Note that each basin is considered independently. Its datum level is its long term minimum water level and not the present value. The maps record the occurrence of high or intermediate levels in each basin within each 1,000 yr period. Millennia characterised throughout by minimal water levels are shown by open triangles. This method underemphasises brief arid periods, but makes best use of isolated ¹⁴C dates, since datable materials are concentrated in lacustrine sediments. For cartographic reasons, basins < 100 km apart have been grouped together.

The reliability of these maps rests ultimately on the 250 internally and stratigraphically consistent ¹⁴C dates on which individual chronologies are based; 79.6% are derived from lake limestone, mollusc and ostracod shells, or other inorganic carbonates, and 12.7% relate to disseminated organic matter or plant macrofossils, and the remainder to charcoal or bone. A large proportion are therefore liable in some degree to the contamination problems besetting continental carbonates⁹. Because of continuing uncertainties about the corrections for variations in ¹⁴C production¹⁰ and initial ¹⁴C/¹²C ratios in lake waters⁹, this study is based on uncorrected dates. This is our main reason for choosing the rather coarse 1,000-yr interval for the maps. We feel that the consistency of the results, in spite of the present spate of new evidence, argues forcibly for the general reliability of the dating framework.

Regional shifts in lacustral conditions

Figure 1 shows the lake-level situation now: the map for the past 1,000 yr is not identical. The distribution of data points is

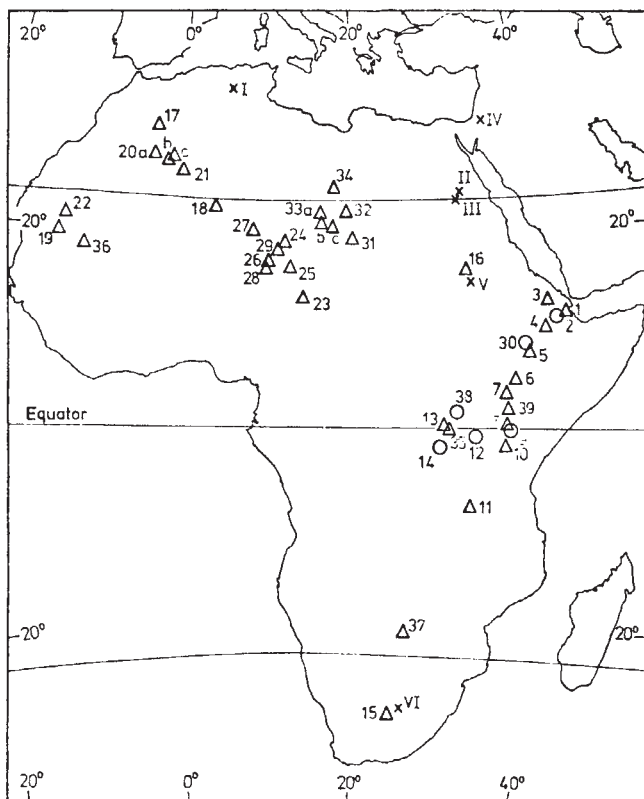


Fig. 1 Present day pattern of African lake levels. Δ , low lake levels; $\circ$, intermediate lake levels; $\bullet$, high lake levels. **Lake Basins:** 1. Assal^{4,14}; 2. Lower Awash (Abbé and adjacent grabens) (Gasse, personal communication; and refs 4,5); 3. Afrera^{5,13}; 4. Middle Awash (Besaka and Gawani) (Williams, personal communication; and ref. 86); 5. Ziway-Shala^{12,37,88}; 6. Stefanie¹²; 7. Rudolf²; 8. Nakuru-Elmenteita^{2,89}; 9. Naivasha^{8,84,89}; 10. Magadi²; 11. Rukwa²; 12. Victoria^{30,90}; 13. Katwe⁴⁰; 14. Kivu³¹; 15. Alexandersfontein²⁵; 16. Jebel Aulia³³; 17. Wadi Saoura, Grand Erg Occidental²²; 18. Hoggar³⁴; 19. Khat Depression²¹; 20. Erg Chech-Touat²² (a) Kadda, (b) Bou Bernous, (c) Bou Ali, Touat; 21. Ahnet-Mouydir Basin²²; 22. Chemchane, Oum Arouaba, Aderg¹⁵; 23. Chad^{6,41}; 24. Kafra, Bilma, Achegour^{6,38,92}; 25. Agadem^{6,92}; 26. Termit Ouest, Kandel-Bouzou^{6,92}; 27. Adrar Bous⁹²; 28. Bouloum Gana⁹²; 29. Fachi^{6,38,92}; 30. Wonchi⁴²; 31. Ounianga Kebir³⁷; 32. Mouskorbé³⁷; 33. Western Tibesti³⁷ (a) Bardagué-Yebbigué, (b) Bégour, Trou au Natron, (c) Tarso Yega; 34. Serir Tibesti³⁹; 35. George⁹⁰; 36. Dhar Tichitt²³; 37. Makarikari (ref. 24 and Grove, unpublished); 38. Albert³¹; 39. Suguta (Bishop, personal communication).

Other localities referred to in text (x)

I Biskra; II Kurkur; III Dungul; IV Dead Sea; V Gezira; VI Florisbad.

patchy, since former lakes are clustered along the African rift valleys and in the closed basins of the Sahara. Suitable basins seem to be lacking within the integrated drainages of the Zaire and Zambezi. But unfortunate gaps exist in present knowledge of southern Africa, the eastern Sahara, and the north African chott region. The effect of volcanic or tectonic activity on rift valley lakes has often been questioned. Careful studies show that lake levels have rarely been significantly disturbed during the past 20,000 yr (refs 11, 12). Even in the particularly unstable Afar Depression, lake-level changes agree remarkably with the rest of tropical Africa^{13,14}.

Figure 1 implies that the present day is a relatively arid period. The few cores available from basins still occupied by lakes (areas 2, 8–9, 12, 14, 38) reveal the occurrence of much greater aridity in intertropical Africa between 21,000 and 12,500 yr b.p.

Figure 2 is a composite map covering the culmination of the last glaciation, 18,000 $\pm$ 3,000 yr b.p. Data for this period are still sparse. We have not plotted any arid intervals which are

unbracketed by ^{14}C dates. A belt of very low lake levels extended from Mauritania across the southern Sahara to the eastern rift. This followed a long period of intermittent high-stand in tropical Africa^{3,15,16} and Arabia (Vita-Finzi, personal communication) which has yielded radiocarbon dates $\geq$ 21,000 yr b.p. (Fig. 6). The level of Lake Abbé began to fall after 23,000 yr b.p., but did not reach its minimum till 17,000 yr b.p. (Gasse, personal communication). Its behaviour closely resembles the most recent data from the Dead Sea¹⁷. Because Fig. 2 simply records the presence of more or less surface water, it should not be used to infer times of transition.

Supporting evidence for widespread tropical aridity at the glacial maximum is summarised by Williams¹⁸. Deflation of desiccated lake sediments and soils readily accounts for the high frequency of freshwater diatoms and opal phytoliths discovered in core K8–57 off the West African coast¹⁹.

The poleward fringes of Africa also experienced lacustral conditions before 21,000 yr b.p., but these persisted intermittently through the next 6,000 yr. Ponding of the southward drainage from the Algerian High Atlas (area 17) created an extensive terrace in which lacustrine levels are dated from > 39,000 to 30,000 and from 24,000 to 14,500 yr b.p. Higher groundwater tables are indicated about 22,900 yr ago in the Western Desert of Egypt²⁰ (localities II and III, Fig. 1) and around 19,000 yr b.p. near Biskra²¹. Significant recharge of aquifers in Tunisia and Algeria continued during the period 21,000–15,000 yr b.p. (refs 22, 23).

Information from southern Africa is still extremely limited. A very large freshwater lake occupied the Makarikari Depression²⁴ at 20,990 yr b.p. (GaK 4310), and lacustral conditions apparently prevailed in South Africa until 16,000 yr ago²⁵. Higher ground water tables from > 32,000 to 14,000 yr b.p. are suggested by widespread spring and cave tufas on both north-west and south-east margins of the Kalahari^{26,27} and by the record provided by spring vent activity at Florsibad^{28,29}.

Maximum aridity

Between 15,000 and 13,000 yr b.p. aridity became even more intense and widespread. Every documented basin in Africa experienced minimum water levels in the fourteenth millennium b.p., with the doubtful exception of Tibesti (Fig. 6). The small fluctuation between 14,500 and 15,000 yr b.p. is probably not real: 4 of the 5 constituent dates come from localities strongly influenced by hot springs (areas 5, 14, 33b). For at least 2,000 yr before 12,500 yr b.p., Lake Victoria, and probably also Lake Albert, ceased to overflow into the Nile^{30,31}. Evaporites were precipitated in Victoria³⁰ and along the White Nile, which "dwindled to a mere seasonal trickle" in the central Sudan³².

Shortly before 12,000 yr b.p. the long arid phase was interrupted in tropical Africa. The Ugandan lakes all regained their overflow levels^{30,31}. Many other basins registered moderate shifts in water balance (Fig. 3). Increasing flooding in both White and Blue Niles gave rise to extensive lakes and swamps in the Gezira (ref. 33; M. A. J. Williams, personal communication, 1975). Downstream from Khartoum, the terrace sequence of the Nile reflects variations in load–discharge relationships, but their interpretation in climatic terms is difficult and some interpretations conflict with this model^{2,32}. Arid conditions still persisted in the north-west Sahara, although small lakes or marshes developed as far north as the Hoggar massif³⁴. These conditions lasted until about 10,500 b.p., when there was a brief drop in lake level in some tropical basins. We suggest that the period 12,500–10,000 yr b.p. is the African counterpart of the late glacial.

Between 10,000 and 9,000 yr b.p., an abrupt change of mode in Arabia (Vita-Finzi, personal communication) and tropical Africa marks the switchover to full interglacial conditions. 18 out of 20 data points suddenly record maximum lake levels. The striking peak at 9,500–9,000 yr b.p. in Fig. 6 is dominated by basal shell dates from thick lacustrine sequences. There is as yet no evidence that the north-west Saharan basins were affected before 8,900 yr b.p., although the South Basin

of the Dead Sea began to overflow around 9,850 yr b.p. (ref. 35). Limited pollen data suggest that conditions in South Africa became warmer and drier³⁶ about this time.

This early Holocene lacustral phase culminated between 8,000 and 9,000 yr b.p., when 31 basins experienced high stands (Fig. 4). The belt of high lake levels extended from Lake Rukwa (8°S 33°E) to the Hoggar (24°N 6°E) and Mauritania (21°N 13°W). Permanent water again existed along the Wadi Saoura. Renewed accretion of lake and swamp deposits took place between the Niles in the Gezira, and the central Saharan massifs supplied significant runoff to surrounding lowlands³⁷.

Lake-level record since 8,000 yr b.p.

Figure 6 shows a more complex behaviour since 8,000 yr b.p. Many basins experienced periods of low levels after 7,500 yr b.p. and did not recover until just after 6,500–6,000 yr b.p. The lesser rise which followed lasted until about 4,800 yr b.p. It was most pronounced in basins fed by groundwater in north-western and central Sahara^{6,15,22,38} and the Afar^{5,13}, in Lake Chad, and in lakes maintained by runoff from the Ethiopian highlands (areas 5–7). Runoff from Tibesti again sustained lakes in the southern desert of Libya³⁹. This rise, which was barely recorded in very marginal basins, marks the end of the Holocene lacustral phase in many areas.

The succeeding 5 millennia have seen a general reduction in surface water storage, punctuated by brief periods of greater runoff which were increasingly restricted in extent. We believe that the peaks in Fig. 6 are more than just random features. The most widespread oscillation prevailed from 3,500 to 3,000 yr b.p. in north-western Sahara, Mauritania, Niger, Chad and the Kenyan–Ethiopian rift (Fig. 5). Runoff from Tibesti into Libya increased although no lakes developed. Since 3,000 yr b.p., desert conditions have once more descended on the central Sahara. Many East African lakes became dry or saline^{2,13,14}. During the past 2 millennia, there have nevertheless been minor moisture fluctuations of $\sim 10^2$ yr, affecting the rift valley, Lake Chad and possibly Mauritania^{12,13,15,40–42}.

Fig. 2 African lake levels for the glacial maximum 18,000 $\pm$ 3,000 yr b.p. Symbols as for Fig. 1.

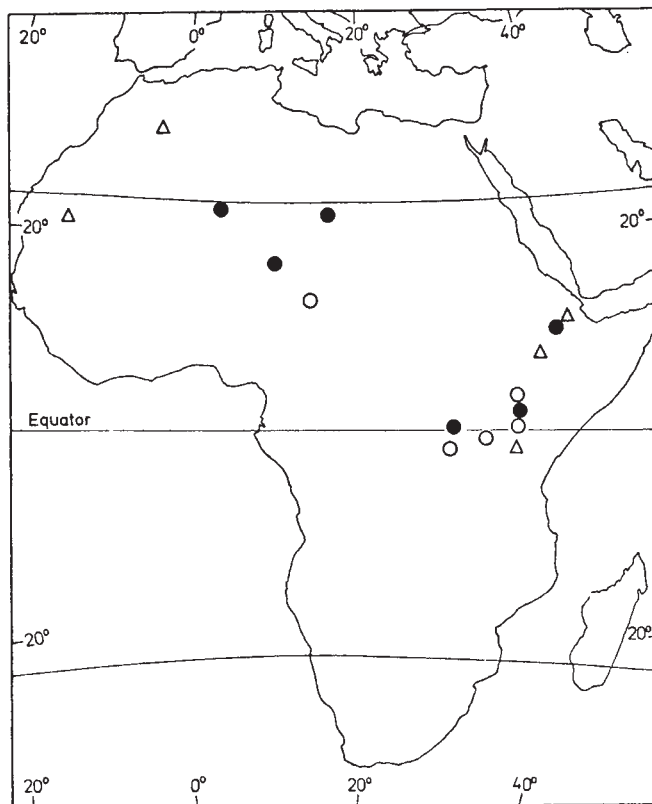
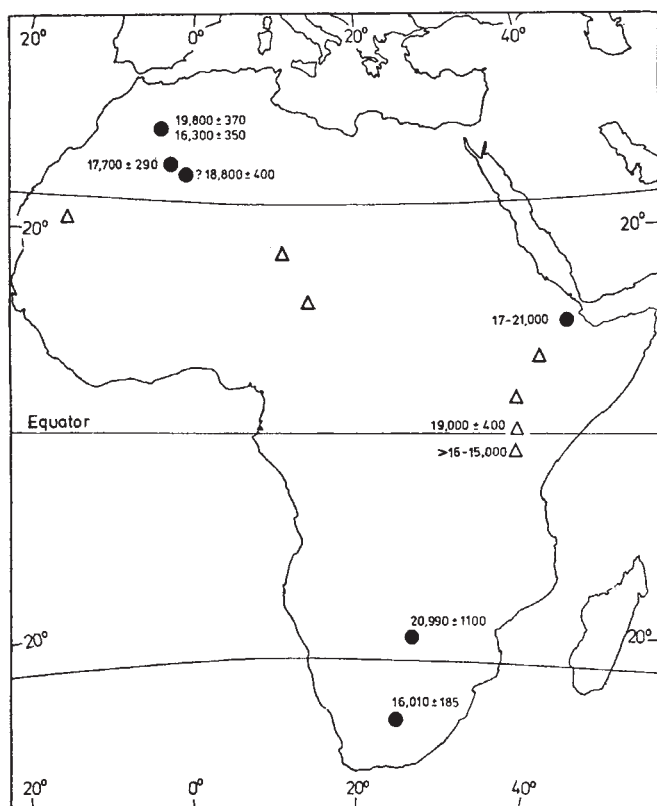


Fig. 3 African lake levels 11,000–12,000 yr b.p. Symbols as for Fig. 1.

These show some coincidence in timing between 2,100–1,500 yr b.p., 1,200–1,000 yr b.p., and after AD 1,600.

Lake levels in the present century have oscillated quite widely on a time scale of 10–30 yr^{12,41,43–45} though the fluctuations show less spatial coherence than in periods 1–3. Most lakes reached high levels between 1870 and 1895, rising again between 1950 and 1965 after several decades of low stand. The extent of Lake Chad peaked in 1965 and contracted dramatically during the Sahel drought⁴¹.

Environmental implication

What are the implications of these profound changes in the abundance of surface moisture? We will focus on four time periods: (1) the glacial mode (21,000–12,500 yr b.p.), Fig. 2; (2) the late glacial (12,500–10,000 yr b.p.), Fig. 3; (3) the interglacial mode (10,000–5,000 yr b.p.), Fig. 4; and (4) the irregular decline from full interglacial conditions (5,000–0 yr b.p.), Figs 1, 5.

During period 1, desert and semi-desert conditions expanded over large areas of tropical Africa. Biogeographic evidence suggests that equatorial lowland forest was drastically reduced in extent and was restricted to refuges in West Africa (ref. 46, Hamilton, personal communication) and eastern Zaire. In the Sudan, desert conditions penetrated up to 450 km southwards into present semi-desert grassland, scrub and thorn savanna areas⁴⁷. The net results were a widespread reduction in effective plant cover and substantial increases in albedo and in infrared emission (10–11 μ m) from sand-covered areas.

There are limited indications of a reverse trend in northern and southern Africa. Mediterranean parkland and scrub may have intermittently invaded the steppelands flanking the Atlas⁴⁸. South African data imply that montane grassveldt displaced semi-desert scrub in the Orange Free State before 17,000 yr b.p. and from 13,200 yr b.p. to 12,600 yr b.p. (refs 28, 36, 48). Events in between are uncertain.

Tropical African river systems responded in a complex way to decreasing plant cover and discharge^{32,33,49}. Arguably,

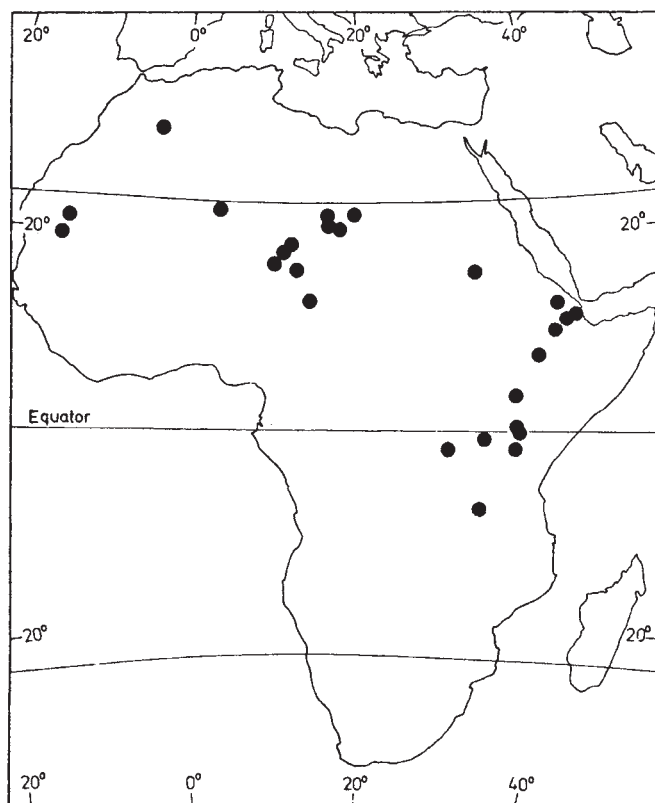


Fig. 4 African lake levels 8,000–9,000 yr b.p. Symbols as for Fig. 1.

the input of solutes to the oceans was greatly reduced, since many major rivers became highly seasonal or were blocked by dunes. The net change in clastic load is still in dispute^{20,32}, and was complicated by sea level lowering. But strong peaks of terrigenous dust in deep-sea cores suggest that aeolian transport of quartz and opalline silica to the Equatorial Atlantic culminated during glacial maxima^{19,50–52}.

The late glacial

Period 2 heralded a general restoration of forest vegetation across central Africa⁴⁶ which was accentuated after 10,000 yr. In the Sahara, semi-desert grassland, scrub and thorn savanna invaded huge expanses of former aeolian sands. Lowland forest species advanced 300–400 km north of their present limits in the Sudan⁴⁷. Formation of lagoonal clays and swamp soils from 12,000 to 4,500 yr b.p. (refs 5, 33) has been of very great agricultural and economic importance.

The interglacial

We infer that during period 3, biological productivity and biomass greatly increased in tropical Africa, and that the surface albedo was substantially diminished by the expansion of open water, marshes and forest. Elephant, giraffe and antelope ranged comparatively unchecked over the Sahara⁵³, which must at times have existed only as relict desert areas isolated by corridors of gallery forest and swamp along major wadis^{47,64}. Overflowing lakes established links between previously separate drainage systems^{12,49}. The significance of these water connections for the spread of aquatic fauna and waterborne diseases is still insufficiently appreciated (see ref. 44). Traces of hippopotamus and crocodile are widespread even in the central Sahara⁵³. Tropical freshwater molluscs, including the vectors of schistosomiasis, are preserved in Holocene deposits from as far north as Tibesti and Hoggar⁵⁶. Malarial mosquitoes probably spread in the same way. The abundance of surface water had a significant role in the diffusion of Late Stone Age fishing cultures and later cattle herders

across the huge region between Mauritania and Kenya^{54,56}. Following this optimum, the decline of the plant cover and the retreat of the big game from the Sahara have been traced by Butzer²⁹ and Wickens⁴⁷.

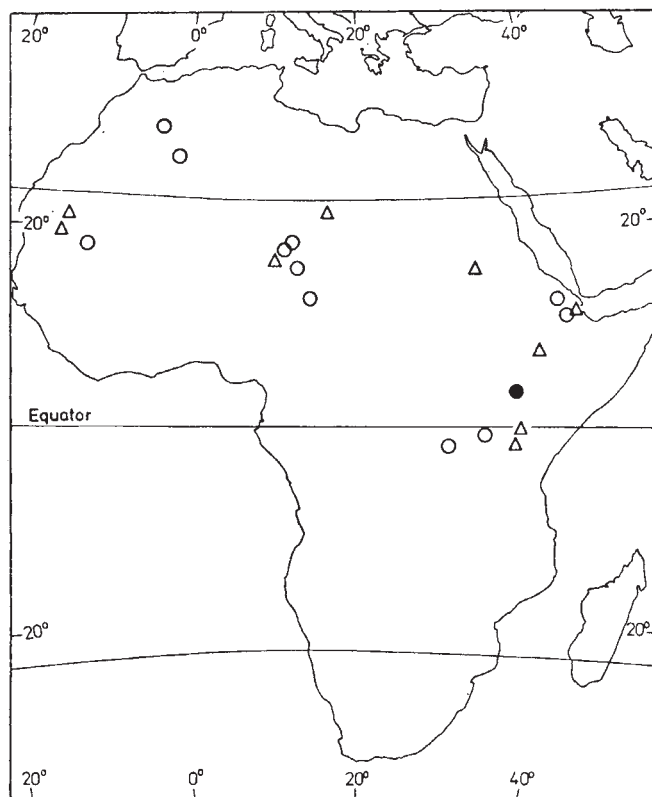
Climatic effects

Since the net radiative cooling of the atmosphere is counteracted mainly by liberation of latent heat⁵⁷, large low-latitude anomalies of precipitation and evaporation are of far reaching climatic significance. Past rainfall can be estimated from the extent of closed lakes, given realistic estimates of temperatures and runoff coefficients. It is assumed, for lack of data, that cloud cover and the seasonality of precipitation have remained constant. For period 1, maximum precipitation estimates for East Africa range from 54 to ~90% of that now, assuming a 3 °C temperature decrease^{30,31,58}. For the Sahel, 15–20% is suggested by environmental indicators^{47,59}. The high levels of mid-latitude lakes could be explained either by increased rainfall or by reduced evaporation caused by lower temperatures, but a 123% increase in rainfall has been computed for Alexandersfontein, based on a 6 °C fall in temperature²⁵. Over Africa as a whole, the net effect is believed to have been a decrease in precipitation. Note, however, that no allowance has been made for the increased vigour of the trade wind circulation⁶¹.

Rainfall estimates for period 3 range from 165% in East Africa (with temperatures equal to or above that of today)³ to 200–400% or more in the Sudan and Mauritania^{47,60}. The East African figures are probably the most representative of intertropical Africa for moisture transport computations.

Palaeoclimatic reconstructions for period 1 imply that (a) the meridional temperature gradient increased in both hemispheres but that the increase was largest in the Northern Hemisphere⁶¹, (b) the vigour of the north-east trades was greatly enhanced^{19,50–52}, (c) Atlantic and Indian Ocean surface temperatures were up to 6–8 °C cooler in areas of upwelling^{62–64}, (d) the Atlantic oceanic polar front was displaced southwards

Fig. 5 African lake levels 3,000–4,000 yr b.p. Symbols as for Fig. 1.



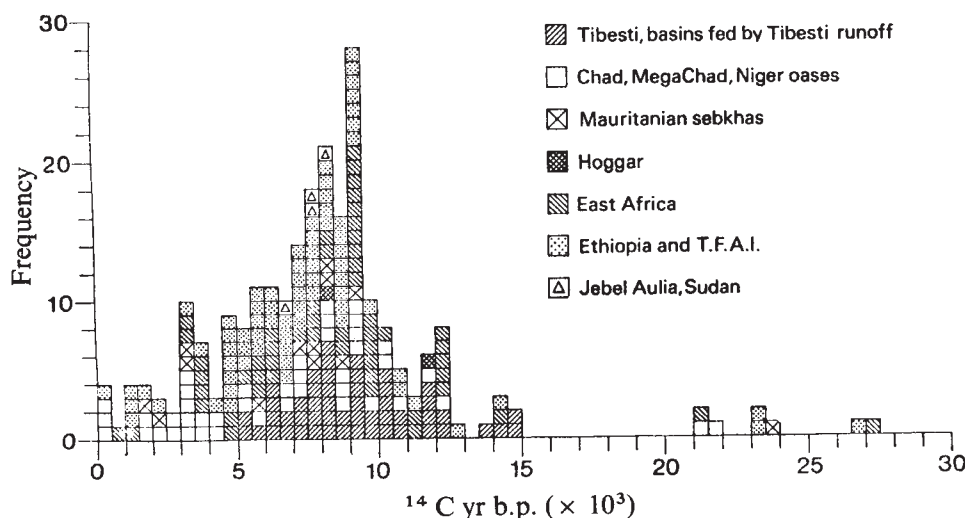


Fig. 6 Histogram of 238 ^{14}C dates from intertropical Africa which record high or intermediate lake levels. Lacustral phases before 21,000 yr b.p. are under-represented because of erosion and burial of their deposits.

to the latitude of Spain⁶², (e) temperatures were 4–8 °C cooler over much of Africa⁶⁵, and (f) global atmospheric moisture transport was substantially reduced^{66,67}. It has been inferred from (a) and (d) that the climatic belts contracted towards the meteorological Equator (g), and from (a) that the mean position of the Intertropical Convergence Zone (ITCZ) lay south of the geographical Equator (h)⁶⁸. Figure 2, coupled with (d) and the evidence for very dry conditions in southern Europe^{69–71}, suggest that the Mediterranean winter rain belt was displaced south over northern Africa, promoting more frequent incursions of cold northerly air. It supports the view that the seasonal migration of the ITCZ and rainfall generation were much reduced, and is compatible with (h) although the evidence from southern Africa is barely sufficient to prove this point (refs 26, 27, 48; and A.T.G., unpublished).

Conclusions

We suggest that the widespread increase in rainfall which culminated between 8,000 and 9,000 yr b.p. was a direct response to increasing ocean temperatures. Figure 6 bears a pronounced similarity to the retreat pattern of polar water in the North Atlantic⁷² including the abortive event seen in period 2. The onset of warming in the western Indian Ocean is dated at 10,000 yr b.p. (refs 64, 73). It is highly significant that this climax of latent heat transport over Africa coincided with the final stagnation and disappearance of the northern hemisphere icesheets^{74,75} and with a major rise in sea level⁷⁶.

Figure 4 implies a significant intensification and greater seasonal displacement of the ITCZ. Data from south of Rukwa are still insufficient to determine whether its mean position lay further north than now⁷⁷. The occurrence of high lake levels in the Holocene in Algeria and Israel^{22,35} suggests a simultaneous deepening of the Mediterranean cyclonic rain belt, although the minimal extent of northern hemisphere ice ~ 5,000 yr ago might allow tropical disturbances to penetrate further north over the Sahara⁷⁸.

Historical lake-level rises seem to have been caused by regional increases in rainfall and runoff⁴³ but are broadly correlated with periods of decreased global temperatures^{41,78,79} and are not a valid analogue for periods 2 and 3. The interplay of the circulations in the two hemispheres was rather different in the earlier periods. But individual years could possibly serve as analogues for periods 1 and 3 although the scale of the rainfall anomalies is smaller. We tentatively suggest the years 1971–72 in the Sahara as models for 21,000–17,000 yr b.p. and 1913 and 1970 for 17,000–13,000 yr b.p. Suitable analogues for period 3 appear to be 1952 and 1957 (refs 78–83). In 1951–52 the 100-mm isohyet was displaced 300 km north in Mauritania⁷⁸, but this advance had been reversed by 1972 (ref. 82).

We conclude that lake-level changes over Africa during the

past 21,000 yr have reflected major shifts in precipitation regimes. The annual isohyet pattern has been displaced for periods of 10^2 – 10^3 yr to the extremes of its recorded distribution (see also ref. 84). These fluctuations have occurred on a sub-continental scale which is directly related to the global dynamics of the atmosphere and oceans. We thank W. W. Bishop, J. M. Bowler, G. L. Isaac, C. Vita-Finzi and M. A. J. Williams for comments. This work has been supported by the NERC, the Royal Society and the Department of Geography, Cambridge.

Received November 13, 1975; accepted April 15, 1976.

- 1 Faure, H., *Mitt. int. Verein. theor. angew. Limnol.*, **17**, 131–46 (1969).
- 2 Butzer, K. W., Isaac, G. L., Richardson, J. L., and Washburn-Kamau, C., *Science*, **175**, 1069–1076 (1972).
- 3 Delibrias, G., Gasse, F., and Rognon, P., *C.r. hebdom. Séanc. Acad. Sci., Paris, Sér. D.*, **277**, 2633–2636 (1973).
- 4 Fontes, J.-C., Moussie, C., Pouchan, P., and Weidmann, M., *C.r. hebdom. Séanc. Acad. Sci., Paris, Sér. D.*, **277**, 1973–1976 (1973).
- 5 Gasse, F., and Rognon, P., *Revue Géogr. phys. Géol. dyn.*, **2**, XV, 405–414 (1973).
- 6 Servant, M., and Servant, S., *Rev. Géogr. phys. Géol. dyn.*, **2**, XII, 63–76, (1970).
- 7 Servant, M., *Séquences continentales et variations climatiques: Evolution du Bassin du Tchad au Cénozoïque supérieur* (ORSTOM, Paris, 1973).
- 8 Bishop, W. W., in *The Late Cenozoic Glacial Ages* (edit. by Turekian, K. K.), 493–527 (Yale University Press, New Haven, 1971).
- 9 Thurber, D. L., in *Calibration of Hominoid Evolution* (edit. by Bishop, W. W., and Miller, J. A.), 1–17 (Scottish Academic, Edinburgh, 1972).
- 10 *Radiocarbon Variations and Absolute Chronology* (edit. by Olsson, I. U.) Proc. 12th Nobel Symp. (Wiley Interscience, New York, 1970).
- 11 Washburn-Kamau, C. K., *Geogr. J.*, **137**, 522–535 (1971).
- 12 Grove, A. T., Street, F. A., and Goudie, A. S., *Geogr. J.*, **141**, 177–202 (1975).
- 13 Gasse, F., Fontes, J.-C., and Rognon, P., *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, **15**, 109–48 (1974).
- 14 Gasse, F., and Stieltjes, L., *Bull. Bur. Rech. géol. min., Sér. 2*, IV, 4, 229–245 (1973).
- 15 Chamard, P. C., *Notes Africaines*, **133**, 1–8 (1972); *Bull. Inst. fond. Afr. noire*, XXXV, 2, 207–243 (1973).
- 16 Butzer, K. W., Brown, F. H., and Thurber, D. L., *Quaternaria*, **11**, 15–30 (1969).
- 17 Kaufman, A., *Geochim. cosmochim. Acta*, **35**, 1269–81 (1971).
- 18 Williams, M. A. J., *Nature*, **253**, 617–618 (1975).
- 19 Parmenter, C., and Folger, D. W., *Science*, **185**, 695–698 (1974).
- 20 Butzer, K. W., and Hansen, C. L., *Desert and River in Nubia* (University of Wisconsin Press, Madison, 1968).
- 21 Williams G. E., *Z. Geomorph. Suppl.*, **10**, 40–63 (1970).
- 22 Conrad, G., *L'évolution continentale post-hercynienne du Sahara algérien* (Sér. géol. 10, CRNS, Paris 1969).
- 23 Thilmans, G., et al., *Bull. Ass. sénégal. Etude Quat. Ouest afr.*, **40**, 5–131 (1974).
- 24 Grove, A. T., *Geogr. J.*, **135**, 191–212 (1969).
- 25 Butzer, K. W., Fock, G. J., Stuckenrath, R., and Zilch, A., *Nature*, **243**, 328–330 (1973).
- 26 Cooke, H. J., *Geogr. J.*, **141**, 430–444 (1975).
- 27 Butzer, K. W., *Curr. Anthropol.*, **15**, 367–382 (1974).
- 28 van Zinderen Bakker, E. M., in *Proc. III Pan Afr. Congr. Prehist.*, 56–67 (edit. Clark, J. D., 1957).
- 29 Butzer, K. W., *Environment and Archaeology* (Methuen, London, 1971).
- 30 Kendall, R. L., *Ecol. Monogr.*, **39**, 121–176 (1969).
- 31 Hecky, R. E., and Degens, E. T., *Late Pleistocene-Holocene chemical stratigraphy and paleolimnology of the rift valley lakes of Central Africa* (Tech. Rpt Woods Hole Oceanogr. Instit., 73–28, 1973).
- 32 Williams, M. A. J., and Adamson, D. A., *Nature*, **248**, 584–586 (1974).
- 33 Williams, M. A. J., and Adamson, D. A., *Geogr. J.*, **139**, 498–508 (1973).
- 34 Delibrias, G., and Dutil, C., *C.r. hebdom. Séanc. Acad. Sci., Paris, Sér. D.*, **262**, 55–58, (1966).
- 35 Neev, D., and Emery, K. O., *The Dead Sea, depositional processes and environments of deposition* (Israel Geol. Surv. Bull. 41, 1967).
- 36 Coetzee, J. A., *Palaeoecol. Afr.*, **3**, 1–146 (1967).
- 37 Geyh, M. A., and Jäkel, D., *Pressdienst Wissenschaft Freien Univ. Berlin*, **5**, 107–117 (1974).
- 38 Faure, H., *Quaternaria*, **8**, 167–76 (1966).
- 39 Pachur, H.-J., *Erde, Berl.*, **106**, 21–46 (1975).
- 40 Morton, W. H., in *Proc. VII Pan Afr. Congr. Prehist. Quat. Stud.* (in the press).
- 41 Maley, J., *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, **14**, 193–227 (1973).
- 42 Smeds, H., *Acta geogr. Helsingf.*, **18**, 1–32 (1964).
- 43 Butzer K. W., *Recent History of an Ethiopian Delta* (Univ. Chicago Dept Geogr. Res. Paper, 136, 1971).

- 44 Beadle, L. C., *The Inland Waters of Tropical Africa* (Longman, London, 1974).
- 45 Degens, E. T., von Herzen, R. P., and Wong, H. K., *Naturwissenschaften*, **58**, 229–241 (1971).
- 46 Hamilton, A., in *East African Vegetation* (edit. by Lind, E. M., and Morrison, M. E. S.), 188–209 (Longmans, London, 1973).
- 47 Wickens, G. E., *Boissiera*, **24**, 43–65 (1975).
- 48 van Zinderen Bakker, E. M., and Butzer, K. W., *Soil Sci.*, **116**, 236–48 (1973).
- 49 Grove, A. T., and Warren, A., *Geogr. J.*, **134**, 194–208, (1968).
- 50 Hays, J. D., and Perruzza, A., *Quat. Res.*, **2**, 355–362 (1972).
- 51 Parkin, D. W., and Shackleton, N. J., *Nature*, **245**, 455–457 (1973).
- 52 Bowles, F. A., *Quat. Res.*, **5**, 225–235 (1975).
- 53 Mauny, R., in *Proc. III PanAfr. Congr. Prehist.*, 102–105 (edit. by Clark, J. D., 1957).
- 54 Gabriel, B., *Palaeoecology of Africa* (in the press).
- 55 Böttcher, U., Ergenzinger, P.-J., Jaekel, S. H., and Kaiser, K., *Z. Geomorph.*, **16**, 182–234 (1972).
- 56 Sutton, J. E. G., *J. Afr. Hist.*, **15**, 527–546 (1974).
- 57 Newell, R. E., Herman, G. F., Gould-Stewart, S., and Tanaka, M., *Nature*, **253**, 33–34 (1975).
- 58 Washbourn, C. K., thesis, Cambridge University, (1967).
- 59 Grove, A. T., *Geogr. J.*, **134**, 194–208 (1958).
- 60 Chamard, P., Guitat, R., and Thilmans, G., *Bull. Inst. fond. Afr. notre, Sér B*, **32**, 688–723 (1970).
- 61 Wilson, A. T., and Hendy, C. H., *Nature*, **234**, 344–345 (1971).
- 62 McIntyre, A., in *Mapping the Atmospheric and Oceanic Circulations and Other Climatic Parameters at the Time of the Last Glacial Maximum about 17,000 Years ago*, 41–47 (University of East Anglia Climate Research Unit Research Publication, 2, 1974).
- 63 McIntyre, A., et al., in *Proc. WMO/IAMAP Symposium on Long-term Climatic Fluctuations*, 75–80 (WMO, 421, Geneva, 1975).
- 64 Vincent, E., *Palaeoecol. Afr.*, **6**, 45–54 (1972).
- 65 van Zinderen Bakker, E. M., and Coetzee, J. A., *Palaeoecol. Afr.*, **7**, 151–181 (1972).
- 66 Kraus, E. B., *Nature*, **245**, 129–133 (1973).
- 67 Newell, R. E., et al., *Nature*, **253**, 33–34 (1975).
- 68 Newell, R. E., *Nature*, **245**, 91–92 (1973).
- 69 Bonatti, E., *Nature*, **209**, 984–985 (1966).
- 70 Wijmstra, T. A., *Acta bot. neerl.*, **18**, 511–527 (1969).
- 71 Florschütz, F., Menendez Amor, J., and Wijmstra, T. A., *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, **10**, 233–264 (1971).
- 72 Ruddiman, W. F., and McIntyre, A., *Quat. Res.*, **3**, 117–130 (1973).
- 73 Olausson, E., *Palaeoecol. Afr.*, **6**, 41–44 (1972).
- 74 Paterson, W. S. B., *Rev. Geophys. Space Phys.*, **10**, 885–917 (1972).
- 75 Zonneveld, J. I. S., *Arct. alp. Res.*, **5**, 223–237 (1973).
- 76 Möner, N.-A., *Geology*, **3**, 109–110 (1975).
- 77 Suzuki, H., *Bull. Dept. Geogr. Univ. Tokyo*, **7**, 1–17 (1975).
- 78 Dorize, L., *Rev. Geogr. phys. Géol. dyn.*, **16**, 393–420 (1974).
- 79 Lamb, H. H., *Geogr. J.*, **132**, 183–212 (1966).
- 80 Grove, A. T., in *Drought in Africa* (edit. by Dalby, D., and Harrison Church, R. J.), 33–45 (Centre for African Studies, SOAS, London, 1973).
- 81 Lamb, H. H., *The current trend of world climate—a report on the early 1970's and a perspective* (University of East Anglia Climate Research Unit Research Publication 3, 1974).
- 82 Wahl, E. W., and Bryson, R. A., *Nature*, **254**, 45–46 (1975).
- 83 Tanaka, M., Weare, B. C., Navato, A. R., and Newell, R. E., *Nature*, **255**, 201–203 (1975).
- 84 Richardson, J. L., and Richardson, A. E., *Ecol. Monogr.*, **42**, 499–534 (1972).
- 85 Gasse, F., *Intern. Rev. Ges. Hydrobiol.*, **58**, 941–64 (1974).
- 86 Page, N., Taieb, M., and Faure, H., in *Proc. VII PanAfr. Congr. Prehist. Quat. Stud.* (in the press).
- 87 Gèze, F., *Bull. geophys. Obs. Addis Ababa* (in the press).
- 88 Haynes, V., and Haas, H., *Radiocarbon*, **16**, 368–380 (1974).
- 89 Richardson, J. L., *Palaeoecol. Afr.*, **6**, 131–136 (1972).
- 90 Stuiver, M., et al., *Radiocarbon*, **2**, 55–56 (1960).
- 91 Trompette, R., and Manguin, E., *Ann. Facult. Sci. Univ. Dakar*, **22**, Sér. Sci. Terre., **2**, 151–162 (1967).
- 92 Faure, H., Manguin, E., and Nydal, R., *Bull. Bur. Rech. géol. min.*, **3**, 41–63 (1963).

Conformation of the lecithin polar group in charged vesicles

H. Hauser & M. C. Phillips

Biosciences Division, Unilever Research Laboratory, Colworth/Welwyn, The Frythe, Welwyn, Hertfordshire, UK

B. A. Levine & R. J. P. Williams

Inorganic Chemistry Laboratory, South Parks Road, Oxford University, Oxford OX1 3QR, UK

The conformation of the polar headgroup of phosphatidylcholine in bilayer vesicles has been investigated using paramagnetic probes. The extended disposition of the polar group when bound to cations is maintained when anions are bound.

A BILAYER structure of phosphatidylethanolamine has been determined from X-ray diffraction studies on single crystals obtained from glacial acetic acid¹. The structure shows the ethanolamine polar group to be approximately parallel to the plane of the bilayer, that is, perpendicular to the long chain fatty acid residues. Although some general structural studies of lecithin layers in water^{2,3} suggest that its phosphorylcholine polar group lies parallel to the long chain residues, that is, perpendicular to the bilayer plane, other studies^{4–6} suggest that the conformation is similar to that found in phosphatidylethanolamine. In all these studies the methods used for structure determination were relatively poor, often involving only one experimental parameter and uncertain theoretical simplifications. We report here a structural study of the polar group of lecithin bound by cations in vesicles exposed to aqueous solution using lanthanide shift and relaxation probes which give us some twelve parameters.

We used paramagnetic probe analysis⁷, involving proton and phosphorus nuclear magnetic resonance (NMR) spectral perturbations of the high resolution spectra of sonicated aqueous lecithin dispersions. The experimental procedure was as follows.

Material and preparation of vesicles

Egg lecithin (grade I), egg lysolecithin (grade I) and glycerophosphorylcholine (as the CdCl₂ complex) were purchased from Lipid Products, South Nutfield. Dilauroyl lecithin was synthesized by acylation of L- α -glycerophosphorylcholine cadmium

chloride using the appropriate acid chloride. The purity of the phospholipids was checked by thin-layer chromatography and that of glycerophosphorylcholine by paper chromatography as before⁸. Single-layer vesicles (~300 Å diameter) were prepared by ultrasonication of a dispersion (typically 1%) of egg lecithin or dilauroyl lecithin in D₂O (99.8%, Norsk Hydro)^{8,9} under nitrogen.

Lanthanide chloride solutions were prepared in D₂O as before¹⁰ to a final nominal pH of 5±0.2. Titrations were carried out by addition of aliquots of the required stock solution to the sonicated lecithin dispersion up to a lecithin to metal molar ratio of 5:1.

Stock solutions of K₃[Fe(CN)₆] and K₃[Cr(CN)₆] in D₂O were prepared after recrystallisation of the compounds (obtained from BDH and Alfa Inorganics, respectively). Aliquots of these solutions were added to the sonicated lecithin dispersion, in the case of Fe(CN)₆³⁻ up to a lecithin to anion molar ratio of 10:1.

Table 1 Assignment of the ¹H resonances of sonicated phosphatidylcholine dispersions in D₂O

–CH ₃ (terminal)	0.88±0.01
(CH ₂) _n	1.27±0.015
CH=CH	5.30±0.015
–CH ₂ –(C=C) ₂	2.82±0.02
–CH ₂ –C=C	2.02±0.02
CH ₂ –C=CO	1.57±0.02
CH ₂ –CO	2.38±0.02
CH ₂ OCO	4.41±0.03
CHOCO	5.32±0.02
CH ₂ OP	4.02±0.02
CH ₂ OP (choline)	4.28±0.01
CH ₂ N	3.69±0.01
N(CH ₃) ₃	3.25±0.01

Chemical shifts are expressed in p.p.m. downfield of TSS.

Table 2 Spin-lattice (T_1) relaxation times (s) of the proton resonances of egg lysolecithin and sonicated egg lecithin dispersions in D_2O

	Lysolecithin	Egg lecithin*
$-CH_3$ (terminal)	0.62 ± 0.03	0.52 ± 0.03
$(CH_2)_n$ (chains)	0.41 ± 0.03	0.38 ± 0.03
$CH=CH$		0.42 ± 0.03
$-CH_2-(C=C)_2$		0.39 ± 0.03
$-CH_2-C=C$		0.39 ± 0.02
CH_2-C-CO	0.29 ± 0.02	
CH_2-CO	0.34 ± 0.02	0.31 ± 0.02
CH_2OCO (glycerol)	$0.38 \pm 0.04^\dagger$	
	0.28 ± 0.03	
CHOCO		
CH_2OP	0.40 ± 0.02	
CH_2OP (choline)	0.35 ± 0.02	0.32 ± 0.02
CH_3N	0.40 ± 0.03	0.31 ± 0.02
$N^+(CH_3)_3$	0.36 ± 0.02	0.32 ± 0.02

*The values given are the average T_1 values for the internal and external groups of the lecithin bilayer as is evident from the determination of spin-lattice relaxation times in the presence of 0.01 M $EuCl_3$ when the internal and external polar group resonances were clearly separated.

† The two protons of the CH_2OCO (glycerol) group are inequivalent and clearly resolved.

1H spectra were recorded on a Bruker 270-MHz instrument operated in the Fourier transform mode and using a superconducting magnet (Oxford Instrument Co.). Trimethylsilylpropane sulphonate (TSS) was used as internal standard. Spin lattice (T_1) relaxation times were obtained by means of a $180^\circ-t-90^\circ$ pulse sequence.

^{31}P spectra were run on a Bruker WH90 spectrometer operating at 36.436 MHz and equipped with heteronuclear decoupling facilities. $(Me_3O)_3PO$ was used as internal standard. All spectra were recorded at an ambient temperature of 26 °C unless otherwise stated.

Determination of binding site in metal complex

A complete assignment of the proton high resolution NMR spectrum of lecithin was required initially. This was based on work reported in the literature together with spin decoupling and nuclear Overhauser experiments¹¹. The confirmed assignments of the observed resonances from groups in the fatty acid chains and the lecithin polar group are given in Table 1. It is worth noting that the $-CH_2OP$ protons (P represents phosphate) of the glycerol are not equivalent, indicating that unlike the CH_2 groups of the choline which are equivalent, the glycerol part of the molecule has a rigid conformation. Studies of line widths and relaxation times have led to this conclusion previously^{9,12}. Spin-lattice relaxation time (T_1) measurements carried out on lysolecithin clearly showed that the T_1 values are shortest in the glycerol part of the headgroup (Table 2).

Before using the paramagnetic lanthanide cations, we added diamagnetic $La(III)$ ions to the lecithin vesicles in aqueous solution to facilitate separation of diamagnetic from paramagnetic effects. As $La(III)$ cations had no perturbing influence on the NMR spectrum, we concluded that the observed shifts with the other rare earth ions were due to paramagnetic perturbations alone. As far as protons are concerned, the shifts due to a metal ion bound to a diester phosphate are known to originate totally from pseudocontact¹³ terms. Two independent procedures showed lanthanide ions to bind only to the phosphate.

(1) The observed ratios of the shifts of the different proton resonances (Table 3) were independent of shift probe^{14,15}. The shift ratio for the ^{31}P resonance relative to the $POCH_2$ proton shifts was, however, strongly dependent on the lanthanide cation. This dependence must be due to contact perturbations (see later) through-bond from the metal.

(2) The measurement of line broadening effects and T_1 values in the presence of $Gd(III)$ as a relaxation probe showed that the nearest protons to this metal were those of the CH_2OP (glycerol) and the $POCH_2$ (choline) groups, but that the phosphorous

atom was considerable nearer than either of these (Table 3).

We conclude that binding of the metal ion to the lecithin molecule occurs at the phosphate group only. Table 3 also shows the relative broadening of various resonances by $Gd(III)$ and this gives an indication of the distance of $Gd(III)$ to the different protons of the polar group and fatty acid chains.

The experiments described above define neither the stoichiometry nor the conformation of the metal binding to the phosphate group. Of interest in the study of the binding of metal ions to a mono- or bilayer structure is the possibility that binding could occur to two centres simultaneously, or even that the observed pseudocontact shifts are a sum comprising contributions from one lanthanide ion bound to a given ligand and another bound on a neighbouring ligand. Two experimental procedures enabled us to examine these possible complications. First, the induced pseudocontact shifts were studied at different degrees of metal binding to the bilayer (that is, with increasing lanthanide concentration up to a lecithin-metal molar ratio of 1:5) and the derived shift ratios were shown to be independent of the number of lanthanide cations bound to the exposed surface of the vesicle in the concentration range used. Second, the shift ratios were determined for vesicles in which the lecithin was diluted by the introduction of an 'inert' molecule, monoolein, to the bilayer. We have found that the shift ratios of the lecithin protons at a given lanthanide cation concentration are independent of monoolein to lecithin molar ratio in the bilayer from zero to about 1.0. We conclude that the observed shift ratios are those for molecules bound to the lanthanide cation and are independent of next near neighbour effects. Third, the stoichiometry of $Ln(III)$:phospholipid at saturation was determined using competition between $Ln(III)$ binding to acetate and to the bilayer. This method showed that one $Ln(III)$ binds to two phospholipid molecules and is therefore effectively chelated on the surface of the bilayer. Strong support for this conclusion was obtained from the shift data with various lanthanides.

As stated above, the shifts for all the proton resonances were measured in the presence of several lanthanide cations— $Pr(III)$, $Nd(III)$, $Eu(III)$, $Ho(III)$, $Tm(III)$ and $Yb(III)$. The magnitude of the shifts for the different ions was in the order predicted by Bleaney¹⁶ and the shift ratios were independent of lanthanide cation even for the usually very anomalous cation, $Tm(III)$ ¹⁷. In the first place whatever the conformation of the headgroup of the lipid, this result shows that the mode of binding in the metal ion-polar group complex is very well defined and independent of cation size. Normally $Tm(III)$ and closely related $Ln(III)$ cations bind to an oxyanion in a different

Table 3 Ratios of the pseudocontact shifts and distances from the probe (r_{ij}) † obtained for the 1H and ^{31}P resonances of lecithin relative to the induced perturbation on the CH_2OP (choline) signal

	R_{ij} *	r_{ij} †
$N^+(CH_3)_3$	0.34 ± 0.04	1.24 ± 0.04
CH_3N	0.55 ± 0.05	1.12 ± 0.04
CH_2OP (choline)	1.00	1.00
^{31}P	3.10 ± 0.40	0.65 ± 0.09
CH_2OP (glycerol)	0.65 ± 0.06	1.03 ± 0.05
CHOCO	0.18 ± 0.08	1.20 ± 0.06
CH_2OCO	0.10 ± 0.06	1.16 ± 0.05
CH_2-CO		1.81 ± 0.12
CH_2-C-CO		1.97 ± 0.14

*Shift ratios, R_{ij} , were independent of the different lanthanide cation used: $Pr(III)$, $Nd(III)$, $Eu(III)$, $Ho(III)$, $Tm(III)$ and $Yb(III)$. The values given relate to studies on egg lecithin, egg lysolecithin and dilauroyl lecithin when the derived shift ratios were invariant within experimental error. No shifts of the protons of the hydrocarbon chain protons were detected in the conditions used.

† Relative relaxation data in the presence of $Gd(III)$ were obtained both by spin-lattice relaxation time (T_1) measurements and line broadening effects (determined by difference spectroscopy). The ratios, r_{ij} , given are those of relative distances from the probe derived from the isotropic (r^{-6}) dependence of the induced relaxation¹⁷.

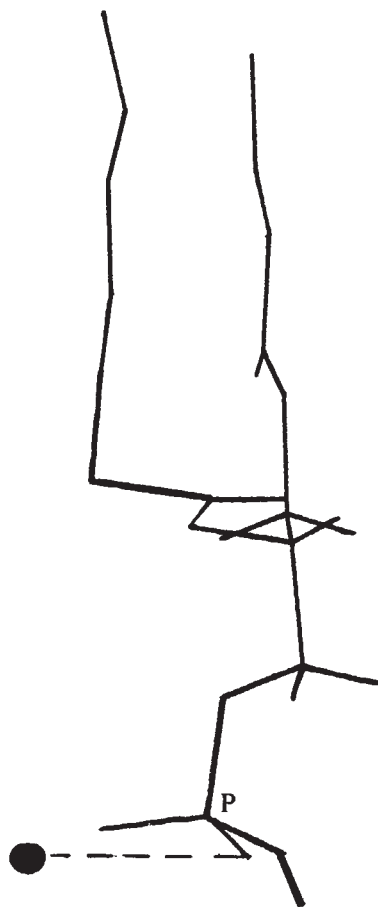


Fig. 1 Metal coordination scheme. The symmetry axis lies along the M-O bond.

manner from the other lanthanides¹⁷. The only previous circumstance in which Tm(III) has been seen to behave in the same way as the other lanthanides in aqueous solution is in binding to ligands which have a very constrained coordination geometry, for example, bidentate (chelate) ligands¹⁸. This observation confirms the deduction from the binding studies that the lecithin-metal complex is a surface 2:1 complex. The constrained conformation is independent of the size of the lanthanide cation, and as we show below, is also maintained when anions are bound. Before turning to anion binding, however, we wish to analyse the conformation of the polar group in detail, using the lanthanide shift probes.

Conformational analysis

The proton shift ratios in the presence of the different Ln(III) cations were found to be independent of the lanthanide cation chosen, and we are therefore justified in the assumption that the site of binding has effective axial symmetry, and that the shifts on the protons are pseudocontact in origin¹³. As shown above, however, the shifts on the ³¹P resonance for the different lanthanides contain both a contact and a pseudocontact term. These were separated by the methods used previously¹⁰. The pseudocontact shift ratio on the phosphorus so obtained (Table 3) merits comment. In comparison with the shift ratios ³¹P:POCH₃ in other phosphate esters, this value is small¹⁰. This is a clear indication that the metal is not bound to the phosphate of lecithin in the same way as was seen for 5'-AMP, for example^{7,13}, when the metal was coordinated to two of the free phosphate oxygens. This is consistent with two observations. First, extrapolation of the measured ³¹P shifts to infinite lanthanide concentration, using a conventional formula for the dependence of binding to the bilayer upon concentration, gave

a value lower than the fully bound shift in simple phosphate ester-lanthanide complexes. Second, the variation of this shift from lanthanide ion to lanthanide ion is not the same as that found in other phosphate esters¹⁰. The magnitude of the ³¹P contact shift due to the lanthanides is also different for the lecithin phosphodiester group from that for all other phosphate esters studied by us^{7,10,13}.

Further demonstration of the degree of difference in the binding mode between lecithin and small model phosphate diesters was obtained from a comparative study of the shifts induced on the protons of glycerophosphorylcholine by lanthanide ions. (Glycerophosphorylcholine has the same chemical formula as the polar group of lecithin.) The shift ratios and the absolute shifts were quite different from those of lecithin and will form the subject of a separate communication.

In the light of these facts a conformational search for the metal ion-lecithin complex could not be based on previous coordination geometries for phosphate ester binding to lanthanides. The results demand an initial computer search for the metal position with respect to the phosphate binding site and the solutions generated must be consistent with the metal ion being in contact with at least one oxygen atom of each of two lecithin phosphate groups, as shown by the ³¹P contact shifts. To proceed with such a metal position search we require a rigid frame portion of the molecule¹⁷, possessing observable nuclear resonances whose spectral perturbations can be used to define the metal position in space. We decided to use the protons of the glycerol group bound to the fatty acid chains since these protons are known to be highly immobilised (see above). Their relative geometry was taken from the structure of phosphatidylethanolamine¹. The diamagnetic spectrum of phosphatidylethanolamine gives a similar proton NMR spectrum as that of lecithin but for the absence of the trimethylammonium resonance. The computational search allowed free rotation about the -CH₂-O-P bonds of the phosphoglycerol unit but spatial overlapping of atoms was forbidden⁷.

The search was made using only one lecithin molecule per metal ion, although there are two lecithin molecules bound per cation. We have shown that the metal complex has axial symmetry. In such circumstances a symmetrical chelating group XA.AX, where A is the binding group (here phosphate), must bind in such a way that groups X and A are related by a rotation about the axis of symmetry or by a centre of inversion at the metal. Thus the metal ion position search can be based on half of the chelate, XA. The solutions are then open to inspection in terms of a constructed dimer which must of course form part of a close-packed bilayer film.

The procedure gave the metal position of Fig. 1 which is consistent with the broadening and T₁ data using Gd(III) as relaxation probe. The position of the metal is not that previously found for the lanthanide complexes of phosphodiester. The conformation of the choline group was then studied using the determined geometry of the metal fatty acid glycerophosphate unit. Free rotation was permitted about the bonds P-O-CH₂-CH₂-N and the conformation giving the best fit to both shift and relaxation data was determined. The centroid of the only family of closely related conformations to fit the experimental data is shown in Fig. 2.

This conformation has the choline approximately perpendicular to the bilayer plane, quite different from the structure found for phosphatidylethanolamine in acetic acid¹. The conformation is also very different from that given by Barsukov *et al.*¹⁹. Relevant internal dimensions and the various torsion angles of the extended group are given in Fig. 2. The study of Yeagle *et al.*⁴ is compatible with this structure although their interpretation is different, and the ¹³C NMR study²⁰ gives data which, for carbon atoms that do not suffer contact shifts, are in agreement with the conformation.

Returning to the 1:2 stoichiometry, we must now place the second lecithin molecule in relation to the first, this position being strictly defined by the rotation about the symmetry axis or centre of inversion (see above).

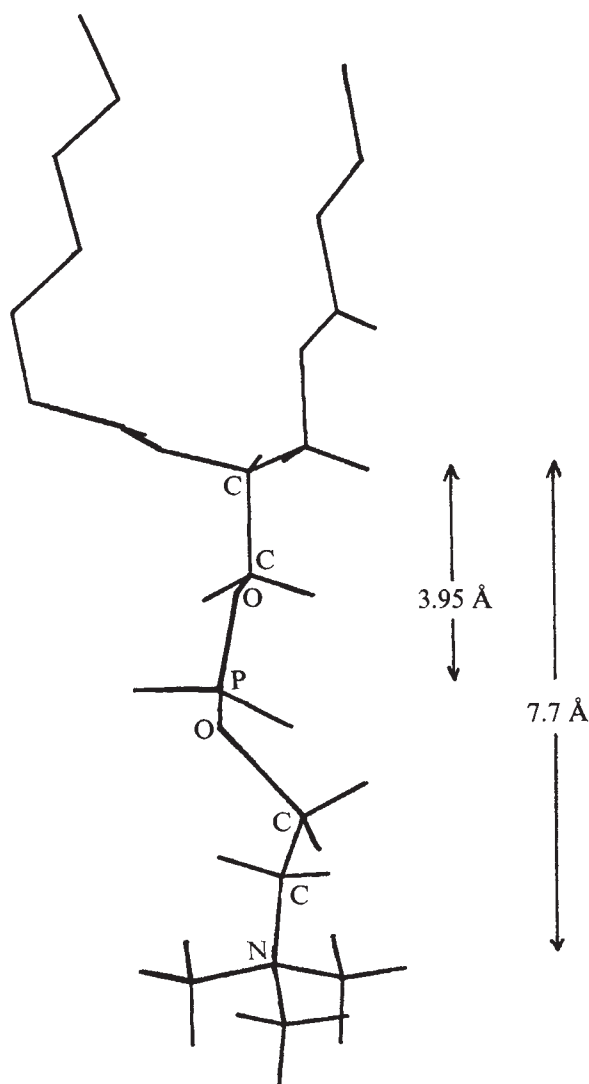


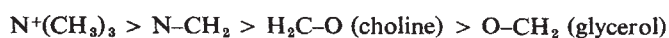
Fig. 2 Conformation of the lecithin polar group. This view is related to Fig. 1 by a rotation of -40° about a vertical axis. The N-C-C-O torsion angle (α_5) is $161 \pm 16^\circ$ while that about the P-O-C-C bond (α_4) is $207 \pm 7^\circ$. The phosphodiester group is in a *gauche-trans* conformation ($\alpha_3 = 201 \pm 6^\circ$, $\alpha_2 = 3 \pm 6^\circ$, $\alpha_1 = 190 \pm 5^\circ$). Torsion angles are defined as in ref. 1.

The axis of symmetry passes along the M-O bond (O of phosphate) so that only one position is permitted for the second molecule which is in the opposite positive quadrant of the shift cone. The resulting bilayer chelate disposes the lecithin molecules about the metal so that they are in the opposite positive regions of the bimodal shift cone function. The fatty acid chains of the lecithin molecules remain in contact with one another and still form a simple bilayer film.

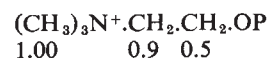
It is interesting to note some internal distances of this chelate formed from the two lecithin molecules. The M-O distance is 2.7 ± 0.1 Å and the separation between the two planes through each N atom (and perpendicular to the bilayer plane) is 11.8 Å.

Paramagnetic anion probes

The conformation of Fig. 2 is open to further experimental investigations through studies of anion binding using probes such as $\text{Cr}(\text{CN})_6^{3-}$ (relaxation) and $\text{Fe}(\text{CN})_6^{3-}$ (shift). Addition of the first of these anions to the aqueous solution of lecithin bilayer vesicles affected the lecithin ^1H NMR spectrum considerably. The order in which resonances were broadened was easily followed by difference spectroscopy and was



The shift ratios with $[\text{Fe}(\text{CN})_6]^{3-}$ were found to be



There was no measurable shift or broadening of other resonances. The data show that the lecithin polar group binds anions in such a manner that the $[\text{Cr}(\text{CN})_6]^{3-}$ is distant from the P-O-CH₂ glycerol group of the lecithin. It is difficult to envisage any conformation of the polar group of lecithin which does not have the $\text{N}^+(\text{CH}_3)_3$ group extended away from the bilayer (as in Fig. 2) but could still give this result. Thus we picture the anions binding considerably further from the fatty acid chains than the cations. This result can be confirmed directly by a consideration of the relative broadening of the protons of the chains induced separately by Gd(III) and $\text{Cr}(\text{CN})_6^{3-}$. It has been found that while Gd(III) causes measurable broadening of the chain protons, no such effects are observed in the presence of $[\text{Cr}(\text{CN})_6]^{3-}$. The final conformation is shown in Fig. 2.

Experiments using lysolecithin permit a detailed confirmation of the above conformation. The proton NMR spectrum of lysolecithin is sufficiently well resolved that we can follow not only the spectral shifts and relaxation times when paramagnetic ions are added, but we can also determine and analyse the coupling constants. Our data show that the conformations of the head groups of lysolecithin and lecithin are the same using both paramagnetic cations and anions to determine distances by relaxation methods and using paramagnetic cations to determine vectors by shift probe methods. As above, data were collected for both protons and phosphorus. Analysis of coupling constants for the choline head-group of lysolecithin gave data very similar to those for choline itself, indicating that the headgroup is in the *trans*¹ configuration. This is the configuration shown in Fig. 2, deduced from shift and relaxation data. The coupling constants are affected only slightly by the addition of metal ions, or increase in temperature.

We do not wish to overstress the significance of the conformation shown in Fig. 2. Our methods^{7,17} give the central solution of a family of conformations all of which, individually and therefore collectively, will fit our data. A considerable population of conformations much outside this family cannot be present, however, for the family is defined by twelve parameters from three methods with different vector functions. Here the family of solutions is relatively small. Finally we cannot exclude motion within the family or to some degree outside it.

Phospholipid polar group conformations

The greatest interest surrounds the disposition of the lecithin polar group (choline) for it is so different from that of ethanolamine. If we consider the two groups, $-\text{N}(\text{CH}_3)_3^+$ and $-\text{N}^+\text{H}_3$, then while the first binds most readily to anions such as RSO_4^- , ClO_4^- and anions high in the chaotropic series, the second polar group binds most readily through hydrogen bonds, that is to RCO_2^- and ROPO_3^{2-} . Sterically $-\text{N}^+\text{H}_3$ is much smaller than $-\text{N}(\text{CH}_3)_3^+$. Presumably together these two factors generate the observed difference in stereochemistry of the two polar groups in their lipid bilayers. We can extend this finding to a consideration of the binding of other molecules to the bilayer surface. Thus phospholipid-protein interactions may show high selectivity. Using the lanthanide probes, it should be possible to map out the position of added molecules lying in the membrane.

The different polar group conformations in the two types of phospholipid may well also give rise to their different behaviour on a macroscopic level. The extended orientation of the phosphatidylcholine polar group could result in a dipolar repulsion between bilayers and the small rate of bilayer fusion²¹. This same property of lecithin may also be the cause of the asymmetric distribution of phospholipids between the inner and outer layer of biological membranes. In erythrocyte membranes,

for example, those phospholipids incorporating the $-N^+(CH_3)_3$ group have been found to be on the outer membrane surface while phosphatidylethanolamines and phosphatidylserines are on the inner membrane surface²². The accumulation of lecithin on the external membrane surface might also inhibit membrane fusion and thus contribute to the stability of cell membranes. The conformational differences between phosphatidylcholine and phosphatidylethanolamine which this work exposes may well account for the differences in their hydration properties²¹.

This work is supported by the MRC.

Note added in proof: The assignment of the proton high resolution NMR spectrum of phosphatidylethanolamine bilayer vesicles has now been completed. Both $-POCH_2$ and $-CH_2N$ appear upfield of their chemical shift positions in the spectrum of lecithin consistent with a different polar group conformation in the two types of phospholipid.

Received October 22, 1975; accepted April 27, 1976.

¹ Hitchcock, P. B., Mason, R., Thomas, K. M., and Shipley, G. G., *Proc. natn. Acad. Sci. U.S.A.*, **71**, 3036–3040 (1974).

- ² Phillips, M. C., Finer, E. G., and Hauser, H., *Biochim. biophys. Acta*, **290**, 397–402 (1972).
³ Jenkasiak, G. L., and Mendible, J. C., *Biochim. biophys. Acta*, **424**, 149–158 (1976).
⁴ Yeagle, P. L., Hutton, W. C., Huang, C.-H., and Martin, R. B., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 3477–3481 (1975).
⁵ Gally, H.-U., Nievesberg, W., and Seelig, J., *Biochemistry*, **14**, 3647–3662 (1975).
⁶ McLaughlin, A. C., Cullis, P. R., Berden, J. A., and Richards, R. E., *J. Magn. Res.*, **20**, 146–165 (1975).
⁷ Barry, C. D., Glasel, J. A., Williams, R. J. P., and Xavier, A. V., *J. molec. Biol.*, **84**, 471–490 (1974).
⁸ Hauser, H., *Biochem. biophys. Res. Commun.*, **45**, 1047–1055 (1971).
⁹ Finer, E. G., Flook, A. G., and Hauser, H., *Biochim. biophys. Acta*, **260**, 49–58 (1972).
¹⁰ Dobson, C. M., Williams, R. J. P., and Xavier, A. V., *J. chem. Soc., Dalton Trans.*, 2662–2664 (1973).
¹¹ Hauser, H., Levine, B. A., Phillips, M. C., and Williams, R. J. P., *Eur. J. Biochem.*, **58**, 133–144 (1975).
¹² Lee, A. G., Birdsall, N. J. M., Levine, Y. K., and Metcalfe, J. C., *Biochim. biophys. Acta*, **255**, 43–56 (1972).
¹³ Barry, C. D., Martin, D. R., Williams, R. J. P., and Xavier, A. V., *J. molec. Biol.*, **84**, 491–502 (1974).
¹⁴ Levine, Y. K., Lee, A. G., Birdsall, N. J. M., Metcalfe, J. C., and Robinson, J. D., *Biochim. biophys. Acta*, **291**, 592–607 (1973).
¹⁵ Huang, C., Sipe, J. P., Chow, S. T., and Martin, R. B., *Proc. natn. Acad. Sci. U.S.A.*, **71**, 359–362 (1974).
¹⁶ Bleaney, B., *J. magn. Res.*, **8**, 91–100 (1972).
¹⁷ Levine, B. A., and Williams, R. J. P., *Proc. R. Soc., A345*, 5–22 (1975).
¹⁸ Bleaney, B., et al., *J. chem. Soc., chem. Commun.*, 791–793 (1972).
¹⁹ Barsukov, L. I., Shapiro, Y. E., Viktorov, A. V., Bystrov, A. F., and Bergelson, A. D., *Dokl. Akad. Nauk. SSSR*, **208**, 717 (1973).
²⁰ Shapiro, Y. E., et al., *Chem. Phys. Lipids*, **14**, 227–232 (1975).
²¹ Hauser, H., in *Water: A Comprehensive Treatise*, **4** (edit. by Franks, F.), 209 (Plenum, New York and London, 1975).
²² Bretcher, M., *Nature new Biol.*, **236**, 11 (1972).

letters to nature

The interplanetary magnetic field structure

THE interplanetary magnetic field is the extension, by the solar wind, of the solar magnetic field¹. The detailed connection between the solar and interplanetary fields depends, however, on the precise manner in which the wind originates in the solar atmosphere and involves a number of complicated dynamical questions which are not yet fully understood². Thus knowledge of the three-dimensional structure of the interplanetary magnetic field is important to our understanding of the solar atmosphere and solar wind flow. Up to now, direct measurements have been confined to the close vicinity of the solar equatorial plane; as a result our ideas about the overall structure of the interplanetary field are speculative. Here we suggest that the cosmic-ray flux provides some indirect information about the morphology of the interplanetary field.

It is well known that the interplanetary magnetic field divides into two or four clearly-defined sectors³, the field in each sector having a constant average sense, towards or away from the Sun. If the fine-scale structure of the photospheric field is ignored, the average field shows large unipolar regions extending over the surface of the Sun. The boundaries between these unipolar regions run in a nearly N-S direction, extending to $\sim 40^\circ$ in latitude^{4,5}. The positions of these boundaries, at the solar equator, correlate with the positions of the sector boundaries of the interplanetary field observed at the orbit of the Earth⁶. For this reason it is generally presumed that the sectors of the interplanetary magnetic field correspond to the large solar unipolar regions and persist to relatively high latitudes⁷, and that the sector boundaries have an orientation nearly perpendicular to the solar equatorial plane.

This conventional view of the interplanetary magnetic



Fig. 1 All positively charged high energy particles, whose centres of gyration fall within one cyclotron radius, R_C , of the surface separating north polar and south polar fields in the solar magnetic cavity, migrate with an average velocity $\langle V_D \rangle$ in the direction of $\nabla \times B$.

field's three-dimensional geometry seems to disagree with measurements showing a seasonal variation in the sector structure^{8–12}. During the six months when the Earth is above the solar equatorial plane those sectors in which the field has the same sense as the north polar magnetic field of the Sun are broader than those sectors in which the field corresponds to the south polar field; and conversely during the six months of the year when the Earth is below the solar equatorial plane. The amplitude of the seasonal bias is large ($\sim 68\%$) and in addition, its phase reverses at the same time that the overall polar magnetic field of the Sun changes sign.

These observations of the interplanetary magnetic field's sector structure seem to suggest that the field in each hemisphere of the solar cavity (north and south) has the same polarity as the average magnetic field at the corresponding solar pole, and that the surface, near the equatorial plane, which separates the two polarity regions is only slightly warped^{2,13,14}. In this picture the observed sector structure results from corotation of the warped surface past the Earth with the angular velocity of the Sun.

Assume for simplicity that, at the orbit of the Earth, the warped surface has the shape of a sinusoid

$$z = z_0 \cos \phi \quad (1)$$

where z is the distance of the surface from the solar equatorial plane and ϕ is solar longitude. The observed 68% amplitude of the seasonal polarity bias implies that $z_0 \approx 0.25$ AU. Thus, on the basis of this unconventional picture of the interplanetary magnetic field one would not expect the sectorized structure to persist beyond $\sim 15^\circ$ of solar colatitude. (By way of contrast solar wind velocity stream structures are known to extend to high latitudes¹⁵.)

Forbush pointed out that the anisotropy which gives rise to the total observed diurnal variation of the high energy cosmic-ray flux is not constant. The diurnal variation apparently consists of two components^{16–19}. The dominant component is constant and corresponds to the corotational anisotropy of cosmic rays. A smaller component, $\sim 25\%$ of the total diurnal variation, corresponds to an anisotropy vector which is aligned approximately with the interplanetary magnetic field and which

oscillates about a zero mean with a full period equal to the 22-yr oscillation time of the solar magnetic field.

We have shown elsewhere^{13,14} that the unconventional picture of the interplanetary magnetic field described above leads in a natural way to the, heretofore unexplained, varying component of the diurnal anisotropy. Briefly, all high energy particles whose centres of gyration fall within one cyclotron radius of the discontinuity separating north and south hemispherical fields undergo a rapid drift $\langle V_D \rangle$ in the direction of $\nabla \times \mathbf{B}$, as illustrated in Fig. 1. Since solar modulation produces a radial gradient of cosmic rays, $dU/dr > 0$, the drift gives rise to an effective source, S , of cosmic rays

$$S = -\langle V_D \rangle_r \frac{dU}{dr}$$

where $\langle V_D \rangle_r$ is the radial component of the average drift. $\langle V_D \rangle$ changes sign from one eleven-year half period of the cycle to the next. S is superimposed on the usual convection-diffusion distribution of cosmic rays. Diffusive reduction of S gives rise to the observed solar-cycle variation in the diurnal anisotropy of high energy cosmic radiation. We know of no other way to account for variation with a 22-yr period²⁰. (Quenby and Hashim²⁰ have discussed the variations in the diurnal anisotropy but gave no natural explanation of the 22-yr cycle.)

Altogether the simple model for the interplanetary magnetic field that we have outlined above provides natural and elementary explanations for two independent and unrelated observational facts. The conventionally accepted field morphology does not lead in any obvious way to either the seasonal bias in the sector structure or to the solar magnetic cycle periodicity of the diurnal variation of energetic cosmic rays.

Since magnetic field lines trace the solar wind flow, comparison of the interplanetary field configuration with the structure of the solar field should provide a valuable constraint on realistic models of the origin of the solar wind. In this regard the discrepancy between the large latitudinal extent of the photospheric sector structure and the apparently small extent of the interplanetary sector structure raises an interesting question. Several resolutions of the discrepancy suggest themselves. For example, the solar wind may originate high in the solar atmosphere where magnetic stresses have simplified the field's geometry. Or, alternatively, it is thought that coronal holes may produce a large fraction of the solar wind. Since large coronal holes are generally present in the polar regions of the Sun, it is possible that most of the solar wind comes from high latitudes and that the latitudinal extent of the sectors is depressed as the polar flow fans out to fill the sphere. A high latitude origin for the solar wind has previously been suggested to account for features of the 11-yr modulation of cosmic rays²¹. (Remote measurements of the solar wind velocity through radio scintillations²² suggest that any fanning out of the solar wind would have to occur $\lesssim 10$ or 15 solar radii from the Sun.) Either of these two possible resolutions of the sector structure discrepancy is consistent with the observation that no more than $\sim 10\%$ of the total photospheric magnetic flux appears in the interplanetary field⁹.

The author is grateful to J. R. Jokipii for valuable discussions. This research was supported, in part, by NASA.

E. H. LEVY

Department of Planetary Sciences
and Lunar and Planetary Laboratory,
University of Arizona,
Tucson, Arizona 85721

Received February 6; accepted April 21, 1976.

- 1 Parker, E. N., *Astrophys. J.*, **128**, 664 (1958).
- 2 Schulz, M., *Astrophys. Space Sci.*, **24**, 371 (1973).
- 3 Wilcox, J. M., and Ness, N. F., *J. geophys. Res.*, **70**, 5793 (1965).
- 4 Wilcox, J. M., and Howard, R., *Solar Phys.*, **5**, 564 (1968).
- 5 Severny, A. B., Wilcox, J. M., Scherrer, P. H., and Colburn, D. S., *Solar Phys.*, **15**, 3 (1970).
- 6 Svalgaard, L., Wilcox, J. M., Scherrer, P. H., and Howard, R., *Stanford Univ. Inst. for Plasma Res.*, No. 639 (1975).

- 7 Gosling, J. T., *Rev. Geophys. Space. Phys.*, **13**, 1053 (1975).
- 8 Rosenberg, R. L., and Coleman, P. J., Jr., *J. geophys. Res.*, **74**, 5611 (1969).
- 9 Rosenberg, R. L., *Solar Phys.*, **15**, 72 (1970).
- 10 Rosenberg, R. L., *J. geophys. Res.*, **80**, 1339 (1975).
- 11 Fairfield, D. H., and Ness, N. F., *J. geophys. Res.*, **79**, 5089 (1974).
- 12 Svalgaard, L., Wilcox, J. M., and Duvall, T. L., *Stanford Inst. for Plasma Res.*, No. 557 (1974).
- 13 Levy, E. H., *J. geophys. Res.* (in the press).
- 14 Levy, E. H., *Proc. 14th Int. Cosmic Ray Conf.*, **4**, 1215 (1975).
- 15 Houminer, A., and Hewish, A., *Planet. Space Sci.*, **20**, 1703 (1972).
- 16 Forbush, S. E., *J. geophys. Res.*, **72**, 4937 (1967).
- 17 Forbush, S. E., *J. geophys. Res.*, **74**, 3451 (1969).
- 18 Forbush, S. E., *J. geophys. Res.*, **78**, 7933 (1973).
- 19 Duggal, S. P., and Pomerantz, M. A., *Proc. 14th Int. Cosmic Ray Conf.*, **4**, 1209 (1975).
- 20 Quenby, J. J., and Hashim, A., *Planet. Space Sci.*, **17**, 1121 (1969).
- 21 Hedgecock, P. C., Quenby, J. J., and Webb, S., *Nature phys. Sci.*, **240**, 173 (1972).
- 22 Coles, W. A., and Maagoe, S., *J. geophys. Res.*, **77**, 5622 (1972).

The microwave spectrum of HNC: identification of U90.7

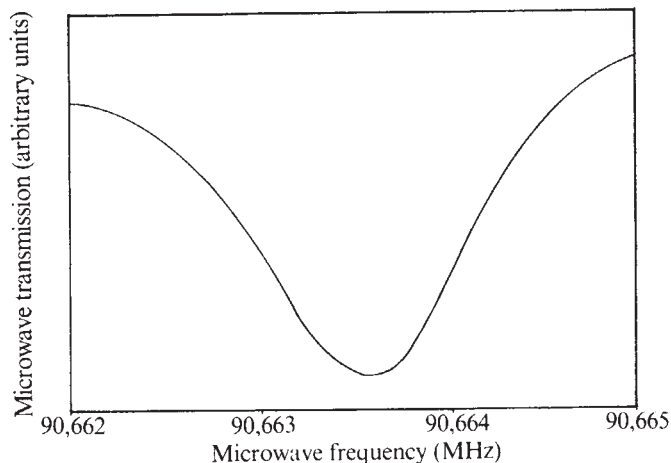
THE detection and characterisation of hydrogen isocyanide, HNC, was for a long time a challenge to science, the first real progress coming from matrix-isolation studies of its infrared spectrum in 1963¹ and it has more recently been detected by its infrared chemiluminescence in certain flames involving atomic nitrogen². In 1971 an interstellar line (U90.7) was attributed by radioastronomers^{3,4} to this molecule, the best estimate of the rest frequency of the line being⁶ 90663.59 ± 0.15 MHz. The possibility that this line represents the $J = 2 \rightarrow 1$ transition of Si C has also been advanced⁵. We have now succeeded in generating HNC in the gas phase and observing its microwave spectrum. The agreement between our observations of the frequency of the $J = 1 \leftarrow 0$ transition and the radioastronomical data render the identification of U90.7 as HNC secure.

In the laboratory the $J = 1 \leftarrow 0$ transition occurs at 90663.59 ± 0.04 MHz for $\text{H}^{14}\text{N}^{12}\text{C}$ (see Fig. 1). This value is in excellent agreement with the astronomical rest frequency for U90.7. We have also observed the corresponding transition for $\text{D}^{14}\text{N}^{12}\text{C}$ at 76305.81 ± 0.04 MHz. From these values the r_0 geometry of HNC emerges: $r_0(\text{HN}) = 98.632$ pm; $r_0(\text{NC}) = 117.256$ pm. These values are in good agreement with r_e values predicted quantum mechanically by a large scale configuration interaction calculation⁷. From Stark effect measurements on both the HNC and DNC lines a dipole moment of

$$(10.2 \pm 0.3) \times 10^{-30} \text{ C m } (3.05 \pm 0.1 D)$$

emerges. The HNC and DNC were generated in sufficient quantities to observe their spectra by heating samples of HNC or DCN to temperatures $\sim 1,000$ K so that the tautomeric equilibrium becomes less unfavourable to the less stable isocyanide. (An earlier attempt to detect HNC at room temperature was unsuccessful⁸.) The lines were observed in a parallel-plate microwave spectrometer, the HCN being 'flash' heated in a series of nozzles to produce broad beams of gas passing trans-

Fig. 1 $J = 1 \leftarrow 0$ transition of HNC. 29 scans averaged. Stark modulation field 700 V cm^{-1} .



versely between the plates and then condensed on a metal plate cooled with liquid nitrogen. Phase-locked klystrons (OKI 80V10A, OKI 90V10A) were used as sources and 33-kHz Stark modulation was used for detection of the signals.

In view of the manner in which we have generated HNC the possibility that the reported transition may be some rovibrational transition of thermally excited HCN must be considered. We eliminate this on the following grounds:

(1) The observed line is narrower than HCN lines. In our conditions the latter are observed as resolved triplets covering ~ 4 MHz, each component having a f.w.h.m. of ~ 800 kHz. The lines here reported have FWHM of 1,300 kHz, their unsymmetrical shape (Fig. 1) suggesting the possibility of hyperfine structure. The observed shape can be quite well simulated by computing the envelope of a multiplet of lines, each of FWHM 800 kHz, with the pattern corresponding to a quadrupole coupling constant $\chi \approx 1.9 \pm 0.3$ MHz. A coupling constant of this magnitude is rather larger than that of other isonitriles (see ref. 9) and is larger than the theoretically predicted constant for HNC ($\chi = 0.9 \pm 0.5$ MHz)⁷. The value is also larger than the upper limit of $\chi \leq 0.83$ MHz that Morris *et al.*⁶ imply from their astronomical observations, and clearly, a higher resolution laboratory spectrum will be needed to obtain a measurement that is free from assumptions concerning either the line shape or the relative intensities of the hyperfine components.

(2) In addition to evidence against this possibility from the line width, the only feasible alternatives for HCN rovibrational transitions in this frequency range have been observed to fall at different frequencies (for example, the $J = 20$ (01¹⁰) l -doublet, observed by Maki and Lide¹⁰ at 93,813.85 MHz) or their frequencies, computed from the formulae presented by Maki and Lide¹⁰, are well removed from those of the lines we report here (the l -doublets for 02²⁰ are at 75.3 GHz and 94.5 GHz for $J = 29$ and 31; for 03³⁰ they are at 72.0, 83.9 and 96.8 GHz for $J = 12, 13, 14$; for 03³⁰ they are at 72.9 and 93.9 GHz for $J = 39$ and 41). From our observations of the 02²⁰, $J = 1 \leftarrow 0$ transition at 89,088 MHz, it is clear that lines derived from more highly excited HCN would in fact have intensities well below our limit of detection in the present apparatus.

The agreement in frequency of the HNC line and that of U90.7 makes the identification of the latter now beyond reasonable doubt. Additional confirmation will come from detection of interstellar DNC at 76.3 GHz and of HN¹³C.

We have also observed a weak signal (of the expected strength corresponding to the natural abundance of ¹³C) from the $J = 1 \leftarrow 0$ absorption of HN¹³C at $87,090.6 \pm 0.2$ MHz, the identification being supported by the facts that the line has a similar shape to that of the main line and that the Stark effect is the same as that displayed by the main line. (We have learned that Dr R. Claude Woods and colleagues at the University of Wisconsin, using a somewhat different laboratory technique, have also observed the main line of HNC at 90,664 MHz.)

We thank Mr J. G. Crofts, Mr R. W. Clarke, Miss M. Haynes and Mr W. G. Matchan for assistance. The work was supported by a grant from the Australian Research Grants Committee.

G. L. BLACKMAN
R. D. BROWN
P. D. GODFREY
H. I. GUNN

Department of Chemistry,
Monash University,
Clayton, Victoria, 3168,
Australia

Received February 26; accepted April 15, 1976.

- ¹ Milligan, D. E., and Jacox, M. E., *J. chem. Phys.*, **39**, 712 (1963); **47**, 278 (1967).
- ² Arrington, C. A., and Ogryzlo, E. A., *J. chem. Phys.*, **63**, 3670 (1975).
- ³ Snyder, L. E., and Buhl, D., *Ann. N.Y. Acad. Sci.*, **194**, 17 (1972).
- ⁴ Zuckerman, B., Morris, M., Palmer, P., and Turner, B. E., *Astrophys. J. Lett.*, **173**, L125 (1972).
- ⁵ Lovas, F. J., *Astrophys. J.*, **193**, 265 (1974).
- ⁶ Morris, M., Zuckerman, B., Turner, B. E., and Palmer, P., *Astrophys. J. Lett.*, **192**, L27 (1974).

- ⁷ Pearson, P. K., Blackman, G. L., Schaefer, H. F., Roos, B., and Wahlgren, U., *Astrophys. J. Lett.*, **184**, L19 (1973).
- ⁸ Blackman, G. L., Brown, R. D., Godfrey, P. D., and Gunn, H. I., *Chem. Phys. Lett.*, **34**, 241 (1975).
- ⁹ Kemp, M. K., Pochan, J. M., and Flygare, W. H., *J. phys. Chem.*, **71**, 765 (1967).
- ¹⁰ Maki, A. G., and Lide, D. R., *J. chem. Phys.*, **47**, 3206 (1967).

Short period fluctuations of the equatorial electrojet

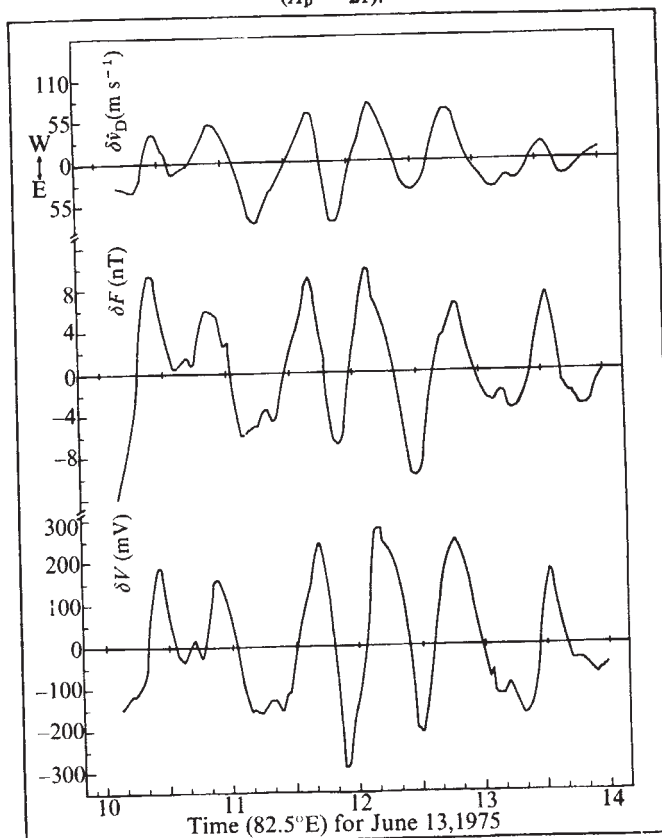
THE signal strength and Doppler frequency spectra of v.h.f. backscatter radar echoes from the equatorial electrojet are being recorded at Thumba near Trivandrum (dip $\sim 1^\circ$ S). The v.h.f. radar operates on 54.95 MHz, with identical arrays of Yagi antennae for transmission and reception; the beam axis is orientated either vertically or obliquely at 30° to the vertical. Analysis of the radar backscatter signals and the magnetograms at Thumba has revealed that quasi-periodic fluctuations of the electrojet on time scales of a few tens of minutes are very frequent. We report here some typical experimental results on this phenomenon.

Radar echoes from only type II irregularities (gradient instabilities) with their characteristically broad Doppler frequency spectra have been analysed. For this type of a broad spectrum, a mean Doppler frequency $\bar{f}_D$ has been defined as

$$\bar{f}_D = \frac{\int_{-\infty}^{\infty} f_D P(f_D) df_D}{\int_{-\infty}^{\infty} P(f_D) df_D}$$

where $P(f_D)$ is the backscattered power per unit frequency at the Doppler frequency f_D . In practice, a summation is used over the usual range of f_D values (-175 Hz to $+175$ Hz). Thus $\bar{f}_D$ is a weighted mean (nearly the same but not identical to the

Fig. 1 Short period fluctuations of the backscatter signal strength (V), the geomagnetic field (F) and the E-W mean drift speed ($\bar{V}_D$) of electrojet irregularities, for June 13, 1975 ($A_p = 21$).



median value) and it shows systematic variations similar to those of the signal strength (and the magnetic field) whereas the Doppler frequency at which the power is maximum shows erratic variations quite often. If we make the usual assumption that the mean drift velocity of type II irregularities is horizontal in the east-west direction, then $\bar{f}_D$ is related to the mean east-west drift speed ($\bar{V}_D$) of the irregularities in the radar-illuminated volume by

$$\bar{f}_D = -(2 \cdot f_0 \bar{V}_D \cos \theta) / c$$

where f_0 is the radar frequency, c is the velocity of light in *vacuo* and θ is the angle of the radar beam axis with respect to the horizontal.

The slow daily variation or 'trend' has been removed in each case by subtracting a running mean from the observed values. A 60-min window is used for obtaining the running mean of the measured parameters which are available at 1-min intervals, except for the Doppler spectra which are recorded at 5-min intervals on most occasions; the 1-min values are then obtained by interpolation. The differences of the 1-min values from the running mean are further smoothed by taking a 6-min running average. The resulting values at 1-min intervals are the smoothed deviations from the mean, representing the fluctuations over time scales of tens of minutes. This procedure is found to bring out quite well the quasi-periodic variations in such a way that they are comparatively free from the longer period 'trend' as well as from the noise-like fluctuations of shorter time scales. (This procedure, however, needs a critical assessment when a quantitative spectral analysis of the quasi-periodic variations is to be made).

Figure 1 shows the quasi-periodic fluctuations in the mean values of (1) the backscatter signal strength (in terms of the receiver detector output voltage), (2) the strength of the total magnetic field (as measured by a Proton Precession Magnetometer), and (3) the mean east-west speed of the type II irregularities as deduced from the mean Doppler frequency of the backscatter signal. The remarkable similarity of the fluctuations in all three parameters and the persistence over a period of 4 h are the two typical features of such fluctuations observed on many days. On this particular day (June 13, 1975), one dominant period of ~ 25 min is present most of the time, but is suppressed part of the time by the interference of another perturbation with a similar period, or a perturbation of a different magnitude and period is dominant for some time. The largest amplitudes of the fluctuations on this day (which was a moderately disturbed day magnetically) are observed around noon and they constitute $\sim 36\%$ of the mean value for the signal strength, $\sim 18\%$ of the daily variation range of the magnetic field, and $\sim 21\%$ of the mean velocity of irregularities. In this figure as well as in the next, the uncertainty in the $\delta \bar{V}_D$ values is variable as it depends on the shape of the spectrum and the power levels measured; for the two cases presented here, the error bars are typically $\pm 10 \text{ m s}^{-1}$. On the other hand, the error bars on δV and δF consist of the measurement errors directly and are $\pm 10 \text{ mV}$ for δV and $\pm 0.7 \text{ nT}$ for δF .

On July 18, 1975, a magnetically quiet day, fluctuations of comparable amplitudes have been observed, as shown in Fig. 2. In this case, the largest amplitudes of the quasiperiodic fluctuations occur near 0800; the peak amplitudes are: 30% of the mean value for the signal strength, 21% of the daily variation range of magnetic field, and 36% of the mean velocity of the irregularities. No single dominant period persists over the interval of ~ 4.5 h; the perturbations of 20–30 min period are present in varying strength.

The main experimental result emerging out of this study is that the electrojet current (and not merely the strength and speed of the electrojet irregularities) undergoes large quasi-periodic fluctuations on many days, including magnetically quiet days. The fluctuations can arise from perturbations (1) in the electron drift speed relative to ions which are not perturbed or (2) in the ion drift speed relative to the electrons which re-

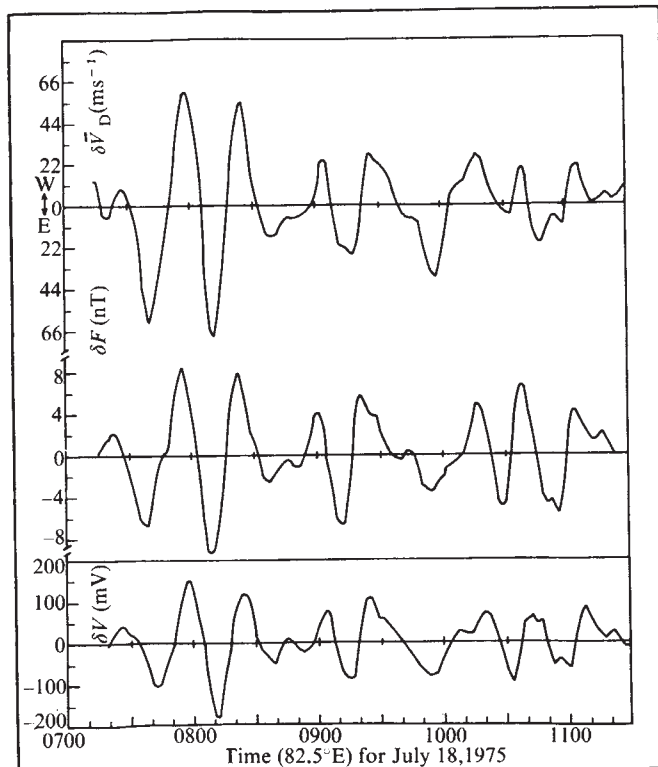


Fig. 2 As in Fig. 1, for July 18, 1975 ($A_p = 6$).

main unperturbed or (3) in both ion and electron drift speeds in such a way that the difference of electron and ion drift speeds undergoes fluctuations.

Theoretical treatments^{1,2} of the motion of ionisation irregularities show that weak irregularities move with the velocity of electrons rather than that of ions. The type II irregularities we have observed are such weak irregularities and the variations in the drift speed $\bar{V}_D$, as observed with the v.h.f. radar, can be interpreted as variations in the speed of the electron drift. An interpretation in terms of variations in the speed of ion drift alone has both observational and theoretical objections: observationally, the measured fluctuations of $\bar{V}_D$ require unrealistically large winds of short periods; theoretically, wind-generated ion motion will give rise to a vertical polarisation field in such a way that the electrons and ions are constrained to move with the same horizontal drift speed³. The physical processes leading to perturbations (1) and (3) above can be understood only through a detailed quantitative theoretical study of the role of mesoscale winds (of gravity-wave origin) in the electrojet region. Such a study is in progress.

In the theoretical treatments^{4–8} of gradient instabilities, the growth time for the type II instabilities is shown to decrease with increasing electron drift speed, but the dependence of the strength of the irregularities on the speed of the electron drift is not clear. The observations do show, however, that the strength of irregularities depends quite sensitively on the electron drift speed (see Figs 1 and 2).

C. A. REDDY
C. V. DEVASIA

Space Physics Division,
Vikram Sarabhai Space Centre,
Trivandrum 695022, India

Received December 30, 1975; accepted March 31, 1976.

- ¹ Clemmow, P. C., Johnson, M. A., and Weekes, K., *Rep. Phys. Soc. Conf. Phys. Ionosph. 1954*, 136 (Phys. Soc., London, 1955).
- ² Kato, S., *J. atmos. terr. Phys.*, **35**, 1073–1082 (1973).
- ³ Kato, S., *J. geophys. Res.*, **78**, 757–762 (1973).
- ⁴ Reid, G. C., *J. geophys. Res.*, **73**, 1627–1640 (1968).
- ⁵ Rogister, A., and D'Angelo, N., *J. geophys. Res.*, **75**, 3879–3887 (1970).
- ⁶ Rogister, A., *J. geophys. Res.*, **77**, 2975–2981 (1972).
- ⁷ Sudan, R. N., Akinrimisi, J., and Farley, D. T., *J. geophys. Res.*, **78**, 240–248 (1973).
- ⁸ Rognlien, T. D., and Weinstock, J., *J. geophys. Res.*, **79**, 4733–4746 (1974).

e.l.f. Hiss modulation at harmonics of the helium gyrofrequency

MANY plasma-wave phenomena involving the proton gyrofrequency have been observed in the ionosphere and magnetosphere but very few involving the helium gyrofrequency. One such is the phenomenon known as helium whistlers which were discovered by Barrington, Belrose and Mather¹ with the Canadian Alouette II satellite. Horita and Friesen² have also reported fine structure at harmonics of the proton and helium gyrofrequency on whistlers detected by the v.l.f. receiver on board the Canadian Isis II satellite. Another paper³, reporting proton gyrofrequency phenomena on auroral hiss, mentioned parenthetically that extra low frequency hiss modulated at harmonics of the helium gyrofrequency has been observed. Here I give further details of the e.l.f. hiss modulation and discuss the region where the modulation may possibly have occurred.

Figure 1 shows the v.l.f. spectrogram made at the Quito station from the Canadian Isis II satellite on April 6, 1971 at 2329.00 to 2329.15 UT; at 2329.10 UT the satellite was at 8.5°S, 74.1°W, and at an altitude of 1,379 km. The superimposed horizontal lines indicate the harmonics of the helium gyrofrequency $f_{He} = 112.5$ Hz (which corresponds to an observed proton gyrofrequency $f_p = 4f_{He} = 450$ Hz). This value was found by examining the spacing of the e.l.f. hiss modulation and is greater than the value of 60.7 Hz for the local helium gyrofrequency in the immediate vicinity of the satellite, suggesting that the source region is at a location or locations beneath the satellite. Note that the horizontal lines generally lie in the gaps of the e.l.f. hiss modulation for frequencies $\geq 2f_p$ and lie on the signals for frequencies $< 2f_p$. The lower frequency boundary of the e.l.f. hiss seems to be close to the local proton gyrofrequency of 243 Hz.

Also shown are whistlers exhibiting the fine structure related to the helium gyrofrequency reported earlier^{2,3}. Note that the whistler to the left at about 2329.05 UT has fine structure such that the horizontal lines pass through the gaps, indicating interaction at a location where the helium gyrofrequency is 112.5 Hz, the same as for the e.l.f. hiss modulation, suggesting a relationship between the two phenomena. The same feature is seen on the whistler to the right at about 2329.10 UT for frequencies > 2 kHz. Below 2 kHz the signal, or perhaps another signal, seems to be displaced a fraction of a second to

the right and has fine structure leading to a helium gyrofrequency of ~ 100 Hz corresponding to $f_p' = 400$ Hz, suggesting a slightly higher interaction region than that for the other whistler. Horizontal lines at these harmonics pass through the gaps in the whistler, as before. This feature suggests attenuation of the whistler at the harmonics of the helium gyrofrequency. For the very faint whistler at about 2329.06 UT, however, the lines pass through the signals, suggesting enhancement at integral values of f_{He} . To be cautious, however, although the fine-structure frequency has been carefully scaled to $\lesssim 1.5\%$ accuracy here, a difference of 2.5% would shift the 20th helium gyrofrequency harmonic line ($5f_p'$) by half a harmonic so that the gaps of the whistler in the box would lie between rather than on the harmonic lines. The phenomenon must, nonetheless, be related to the helium gyrofrequency, and for those examples with a large number of harmonics the very high order harmonic frequencies can be more accurately determined.

Other examples of the e.l.f. hiss modulation gave values for the observed helium gyrofrequency which were approximately equal to and also less than the local value in the immediate vicinity of the satellite, depending on whether the satellite was somewhat near the plasmopause or outside the plasmasphere respectively. These observations are consistent with the statement by Taylor and Shawhan⁴ that the propagation of e.l.f. hiss is downgoing outside the plasmopause and upgoing inside.

Since the e.l.f. hiss modulation provides a signature for the source region or regions where the modulation took place, it provides a means for studying the possible generation or interaction region of the e.l.f. hiss depending on whether the modulation took place at the same time as or after the e.l.f. generation. A preliminary calculation for the altitudes of the source regions, assuming wave propagation along the magnetic field of the Earth, gave $\sim 2,300$ km outside the plasmasphere, 500 km inside the plasmasphere and 1,400 km (the Isis II satellite altitude) near the plasmopause. Obtaining the observed gyrofrequency used to calculate the location of the generation or interaction region can be very complex since at times three or more components may be seen on the v.l.f. spectrograms, indicating a superposition of many e.l.f. waves from a number of source locations.

Work is in progress to analyse more examples of e.l.f. hiss modulation. The results of the analyses and more complete details of the items mentioned above will be published elsewhere. Further studies of the phenomenon, especially with data from satellites at other altitudes than the nearly constant Isis II satellite altitude of 1,400 km, may provide further clues as to the generation mechanism for e.l.f. hiss, the subject of some discussion in very recent papers⁵⁻⁹.

I thank Mr Lane Friesen for assistance in the preparation of Fig. 1, Drs R. E. Barrington and F. H. Palmer for the microfilmed Isis II v.l.f. spectrograms, the World Data Center A for Rockets and Satellites at the Goddard Space Flight Center for supplying microfilmed world maps for the Isis II satellite and the NRC of Canada and the University of Victoria for support.

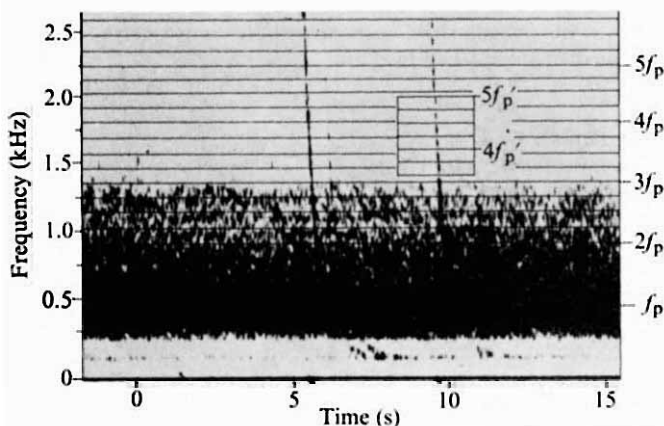
R. E. HORITA

Department of Physics,
University of Victoria,
Victoria, British Columbia,
Canada

Received March 15; accepted April 15, 1976.

- Barrington, R. E., Belrose, J. S., and Mather, W. E., *Nature*, **210**, 80-81 (1966).
- Horita, R. E., and Friesen, Lane, *J. atmos. terr. Phys.*, **37**, 1497-1500 (1975).
- Horita, R. E., Friesen, Lane, and Chan, A. W. Y., *J. atmos. terr. Phys.* (in press).
- Taylor, W. W. L., and Shawhan, S. D., *J. geophys. Res.*, **79**, 105-117 (1974).
- Russell, C. T., McPherron, R. L., and Coleman, P. J., Jr, *Space Sci. Rev.*, **12**, 810-856 (1972).
- Thorne, R. M., Smith, E. J., Burton, R. K., and Holzer, R. E., *J. geophys. Res.*, **78**, 1581-1596 (1973).
- Etcheto, J., Gendrin, R., Solomon, J., and Roux, A., *J. geophys. Res.*, **78**, 8150-8166 (1973).
- Parady, B. K., et al., *J. geophys. Res.*, **80**, 2183-2198 (1975).
- Thorne, R. M., and Barfield, J. N., *Geophys. Res. Lett.*, **3**, 29-32 (1976).

Fig. 1 e.l.f. hiss exhibiting modulation related to the helium gyrofrequency. The fine structure leads to a helium gyrofrequency $f_{He} = 0.25f_p = 112.5$ Hz, which is greater than the local value of 60.7 Hz in the immediate vicinity of the satellite; the superimposed horizontal lines are at harmonics of f_{He} . Also shown are whistlers which may be related to the e.l.f. hiss.



Thermoelastic instability in lubricated sliding between solid surfaces

IN the past few years, it has been recognised that, in certain conditions, thermoelastic instabilities occur in dry sliding contact between two solids¹⁻⁴. We here report observations of the same phenomenon in the presence of a liquid lubricant which supports analytical predictions⁵ of such an effect. Thermoelastic instability is caused by frictional heating and thermal deformation of the surfaces in contact, and alters the original contact configuration. Even if the surfaces are perfectly flat at the outset, instability will lead to a departure from flatness, ultimately resulting in patches of high pressure in the lubricant, with surrounding regions of low pressure and relative parting of the surfaces. This instability can lead to a malfunction of, for example, mechanical face seals.

If the instability leads to surface deformations, stability implies the suppression of any disturbance of the initial surface shape. The two regimes are demarcated by a threshold on which neutral stability prevails; on this threshold, a disturbance on either surface will be preserved without growth or decay. Several such thresholds exist in the case of a good thermal conductor sliding on an insulator⁵. Figure 1 shows an enclosed region, defined by the sliding speed and the film thickness, within which instability is predicted. The bottom edge of this region is the threshold of principal interest and the corresponding critical sliding speed is given by

$$U_{crit} = \bar{h} \kappa (K_M / \mu \alpha_M)^{1/2}$$

where $\bar{h}$ is the nominal film thickness, K_M and α_M the coefficients of thermal conductivity and expansion, respectively, of the conductor, μ the viscosity of the oil and κ is a wave number of the surface disturbance, being given by $2\pi/\lambda$, where λ is wavelength. It is also predicted that this instability will occur for only a range of $\bar{h}$, as shown in Fig. 1. These results were obtained from an idealised analysis where we tried to determine whether or not a small perturbation would grow on a surface that, initially, was perfectly flat. The analysis was not extended to the nonlinear effects that would eventually result from such a

Fig. 1 Domain of instability for aluminium sliding on glass. The solid lines represent analytically predicted thresholds of ref. 5 for $\kappa = 0.24 \text{ cm}^{-1}$; the dashed line corresponds to $\kappa = 0.48 \text{ cm}^{-1}$.

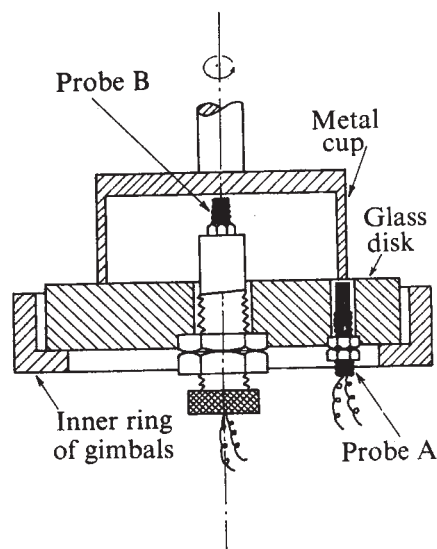
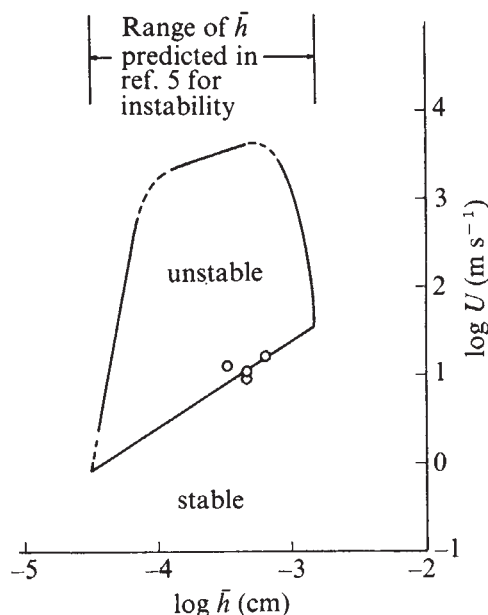


Fig. 2 Schematic diagram of experimental set-up, showing the location of the two measuring probes.

growth. But the experiments reported here show that such instabilities can occur in practice.

To demonstrate the effect, a circular metal cup with a flat, thin lip is rotated over a flat Pyrex disk (Fig. 2). The cup is rigidly held in a vertical spindle which has freedom of motion along its axis. The glass disk is flexibly mounted in gimbals. Placed in two holes in the glass, and rigidly affixed to it, are two eddy-current proximity probes. One of these (probe A) records the change of the distance between the lip of the cup, as it turns, and the tip of the probe. The other (probe B), located on the cup axis, records axial vibrations or bouncing of the cup as a whole. No attempt is made to control the film thickness, it being controlled by the operating conditions. When probe B indicates no bouncing of the cup, the time dependent output of probe A may be taken as the measurement of the contour of the surface passing over it. The wavelength λ will be an integral submultiple of the cup circumference.

At the outset, the surfaces of the cup and the glass are ground and polished to a mirror finish, with a waviness of $\sim 10^{-4} \text{ cm}$ peak to peak. Oil is held in a shallow pool on the glass. The two probes are connected to a dual beam oscilloscope and each is also connected, in parallel, to a high impedance voltmeter for visual reading of displacements. As the cup begins to turn, the oil temperature rises, causing thermal expansion of probe A and, consequently, a slow drift in its reading, without affecting its dynamic response to the passing metal surface. For this reason, an initial running in period, at low speed, is allowed until the d.c. voltmeter reading for probe A stabilises. The rotational speed is then gradually increased in steps and time is allowed at each step for stabilisation of the d.c. voltmeter reading for probe A.

The zero shift with temperature makes it difficult to measure directly the time-average thickness of the liquid film. A technique has been adopted in which the d.c. voltmeter reading is allowed to stabilise, and then the rotation of the cup is suddenly arrested causing it to drop and rest on the glass. Voltages (d.c.) before and after this stoppage are noted and simultaneous photographs are taken of the oscilloscope traces. To this d.c. voltage drop is added the mean of the fluctuating (a.c.) component as indicated by the oscillogram for probe A before stoppage, to obtain $\bar{h}$ as defined in ref. 5. This procedure is carried out at several rotational speeds.

In Fig. 3 the upper trace corresponds to probe B and indicates d.c. voltages, the lower trace indicates the fluctuating (a.c.) component of the voltages from probe A. A tendency towards parting of the cup and the glass is represented by a rise in the

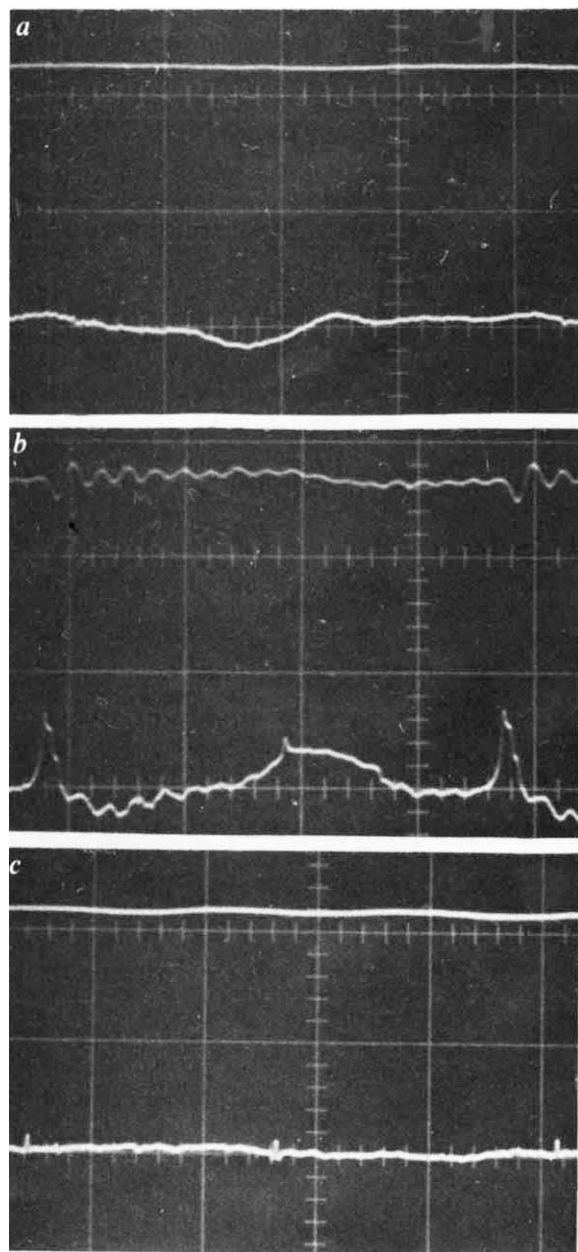


Fig. 3 Oscillograms at three different speeds. Upper trace corresponds to probe B, lower trace to probe A. One small vertical division corresponds to 0.004 mm. Each trace shows a little more than one revolution. *a*, Surface profile at 10.7 m s^{-1} speed, subcritical. *b*, Thermoelastically deformed surface at 13 m s^{-1} sliding speed, supercritical. *c*, Surface profile at 4.6 m s^{-1} , on slowing down from the deformed state; large deformations have disappeared; indicating that they were thermoclastically generated at the higher speed.

top trace and a drop in the bottom one, both to the same scale. Figure 3a shows the surface profile at a sliding speed a little below the critical. This profile was seen to change with time, with small thermally-generated bumps growing and moving on the surface, but with the surface wave amplitude remaining approximately constant. This represents a condition close to neutral stability, and has been observed repeatedly.

Figure 3b exhibits a dramatic deformation of the face at a somewhat higher speed, with large thermal asperities forming on the surface. This is taken as experimental evidence of the onset of the predicted instability. As a large asperity passes over probe A and dips into the slight depression above the probe, it sets off a vibration that is clearly indicated by the upper trace.

This vibration is damped out until the next asperity passes over the probe and triggers another vibration. Figure 3c is taken at a lower speed after brief operation in the deformed condition of Fig. 3b. The asperities have all but disappeared, supporting the argument that they are not the manifestation of a permanent surface damage but are of thermal origin. Observation of the specimen after the experiment suggests that the asperities in Fig. 3b have broken through the oil film and are rubbing directly against the glass.

The qualitative aspects of these observations are repeatable, although the details of the deformed state change somewhat from specimen to specimen. For an aluminium cup of 7.62-cm diameter and 0.15-cm lip thickness, and with oil similar to SAE 10, the sliding speeds and film thicknesses corresponding to the onset of the deformed state are shown in Fig. 1 as data points. The solid line is the theoretically predicted threshold of instability for a λ of one full circumference ($\kappa = 0.24 \text{ cm}^{-1}$), the dashed line for a κ of one-half the circumference ($\kappa = 0.48 \text{ cm}^{-1}$).

This work was supported primarily by the US Office of Naval Research.

B. N. BANERJEE
R. A. BURTON

Department of Mechanical Engineering and
Astronautical Sciences,
Northwestern University,
Evanston, Illinois 60201

Received January 12; accepted April 22, 1976.

¹ Barber, J. R., *Proc. R. Soc.*, A312, 381–391 (1969).

² Dow, T. A., thesis, Northwestern University, Evanston, Illinois, USA (1972).

³ Burton, R. A., Nerlikar, V., and Kilaparti, S. R., *Wear*, 24, No. 2, 177–188 (1973).

⁴ Burton, R. A., and Nerlikar, V., *Wear*, 27, No. 2, 195–207 (1974).

⁵ Banerjee, B. N., and Burton, R. A., *J. Lubric. Technol.* (in the press).

Linear dichroism of cations and anions in micellar solutions

WE here report a linear dichroism (LD) exhibited by isotropic chromophores in a condensed anisotropic medium and caused by anisotropic effective local fields. We shall call the effect Wiener dichroism (WLD) when the anisotropic reaction field arises from the presence of two (or more) different dielectric domains (phases in our notation) which are internally isotropic but which are separated by surfaces with non-spherical geometry. With a coloured cubic anion, like $\text{Co}(\text{NO}_2)_6^{3-}$, in the aqueous phase of a 20% solution of cetyltrimethylammoniumbromide (CTAB), known to contain rod-shaped micelles, a positive WLD is observed, while a cation, like $\text{Co}(\text{NH}_3)_6^{3+}$ yields negative WLD, when the micelles are oriented. This difference between cations and anions arises from the different average locations of the ions with respect to the cylindrical counter-ion layer surrounding the micelle. This is the first time a dichroism has been observed caused exclusively by local fields.

In the applications of linear dichroism spectroscopy^{1–6} it has been accepted that the observed absorptions $A_{\parallel}$, $A_{\perp}$ in the macrosystem (for example a crystal) are directly related to transition probabilities in the microsystem (molecule) being proportional to the squares of the scalar products between the radiation electric field vector and the transition moments (sometimes after correction for effects on the polarisation of the field if the axes of the indicatrix or the linear dichroism ellipsoid do not coincide with the laboratory directions⁷). Though it seems intuitively clear that with anisotropic permittivity and hence different effective fields in different directions, $A_{\parallel}$ may differ from $A_{\perp}$, this has never been experimentally verified, even by molecular absorption techniques, and it has been simply assumed that the 'molecular dichroism' is proportional to the observed macroscopic dichroism.

We consider the case of elongated oriented particles (molecules, domains or crystallites) immersed in an isotropic medium. Wiener and other authors^{8–10} estimated the macroscopic static permittivity ϵ in terms of the permittivity of the particle, ϵ_{2i} ($i = \parallel$ to the orientation axis and $\perp$, respectively), and of ϵ_1 ,

the permittivity of the surrounding 'phase'. If the particle is approximated by an ellipsoid of revolution, the applied uniform field, E_0 , is modified inside the ellipsoid according to simple electrostatic theory, equation (1a), where L , the depolarising coefficient, depends on the geometry of the particle⁹. If the particles define the volume fraction f , the apparent average field, $E = (1-f)E_0 + fE_i$, yields ϵ_i of the system: equation (1b).

$$E_i = E_0/[1 + L(\epsilon_{2i} - \epsilon_1)/\epsilon_1] \quad (1a)$$

$$\epsilon_i = \epsilon_1 + f(\epsilon_{2i} - \epsilon_1)/[1 + (1-f)L(\epsilon_{2i} - \epsilon_1)/\epsilon_1] \quad (1b)$$

This or similar expressions have been used to explain 'form birefringence' in oriented polymer solutions¹¹ and in crystals, for example, by the change in birefringence when water in a haemoglobin crystal was replaced by a salt solution (that is ϵ_1 changed) Perutz estimated the shape of the protein molecule¹², but a complex permittivity was not considered. By inserting $\epsilon_{2i} = (n_{2i} + i k_{2i})^2$, $\epsilon_1 = (n_1 + i k_1)^2$, eq. (1) may be solved with respect to the macroscopic refractive indices n_i , n_1 and absorption coefficients k_i , k_1 in terms of the particle coefficients, k_{2i} , k_{21} , the absorption of the surrounding medium, k_1 , and the corresponding refractive indices, n_{2i} , n_{21} , n_1 .

The molecular linear dichroism $k_{2i} - k_{21}$ can thus be correlated to the observed dichroism, $k_i - k_1$ (relationship between k and the decadic absorption coefficient in eq. (2)). Wiener arrived at a similar conclusion 1912 but (probably because of the great length of his paper) his work has unfortunately remained largely unknown⁸. What happens if the particle is isotropic, that is $\epsilon_{2i} = \epsilon_{21}$? Eq. (1) then yields eq. (2), a linear dichroism which has its origin purely in the non-spherical geometry of the surfaces separating the dielectric phases, and which we call Wiener dichroism, WLD, in memory of its predictor (A is decadic absorbance, d pathlength).

$$WLD = A_i - A_1$$

$$A_i = k_i(4\pi vd/\log_e 10)$$

$$k_i = \frac{(1-f)k_1\{n_1^4 + L_i(n_2^2 - n_1^2)[2n_1^2 + f(n_2^2 - n_1^2)] + L_i^2(1-f)(n_2^2 - n_1^2)^2\} + fn_1^2 n_2 k_2}{\{n_1^2 + [f + (1-f)L_i](n_2^2 - n_1^2)\}^{1/2} [n_1^2 + (1-f)L_i(n_2^2 - n_1^2)]^{3/2}} \quad (2)$$

$L_{||} = L_{\perp} = 1/3$ for a sphere, $L_{||} = 0$, $L_{\perp} = 1/2$ for an infinite cylinder

If only the surrounding of the infinite cylinders absorbs ($k_2 = 0$) equation (2) may be approximated to $WLD/A = -2f(1-f)[(n_2/n_1) - 1]$, where A is the unpolarised absorbance. This corresponds to the case of CTAB (20%) in an absorbing aqueous solution. Inserting $f = 0.20$, $n_1 = 1.33$ (water), $n_2 = 1.40$ (CTAB) we get $LD/A = -0.017$ (500–600 nm), which in sign

Fig. 1 LD/A according to equation (2), $f = 0.20$; a, $n_2/n_1 = 1.4/1.33$; b, $n_1 = 1.37$ ($\square$), 1.33 ($\circ$) and 1.29 ($\triangle$).

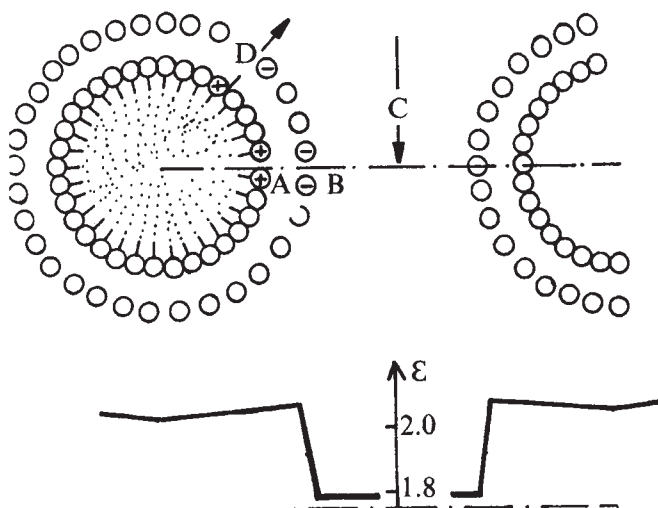
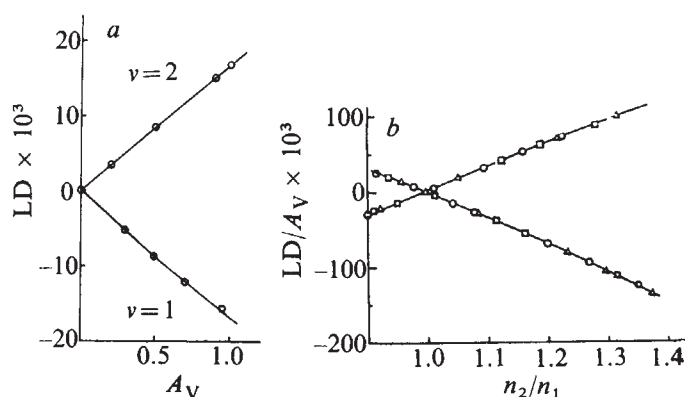


Fig. 2 Sketched structure in plane perpendicular to orientation direction of CTA⁺ micelle cylinders surrounded by Br⁻ cylinders, below schematic variation in permittivity.

and magnitude agrees with the observed effects of some isotropic chromophoric cations (Table 1) when the system is subjected to a shear gradient $> 1,000 \text{ s}^{-1}$ in a Couette cell¹³. The LD spectrum has the same shape as the absorbance spectrum, that is LD/A is approximately constant over the spectrum (Table 1).

Further, LD/A appears independent of the chromophore concentration, which is expected since the influence of the latter on n_1 is negligible at these concentrations.

The independency of LD/A of wavelength is evidence for completely negligible orientation dichroism ($k_{1||} = k_{1\perp}$): thus if, for example, the trigonal $\text{Fe}(\text{dipy})_3^{3+}$ had not been randomly orientated, positive and negative LD contributions should have

appeared corresponding to the existence of mutually perpendicularly polarised transitions. Thus the non-cubic complexes also behave isotropically.

Finally an 'anomalous' behaviour of anions (Table 1), exhibiting positive LD, suggests an interesting possibility of using WLD for solution structural studies: a different LD/A implies a different average location with respect to the micelle cylinder (Fig. 1). The sign reversal may thus be explained by the sketched permittivity profile, which is expected if the bromide ions form a compact coaxial cylindrical mantle. The anions are preferably found in, or inside, the mantle (A), attracted by the solid positive charge, while the cations are outside (or in contact with) the bromide mantle (B). Figure 2 shows effects of different refractive indices in the cases of absorber inside and outside the mantle surface, according to equation (2). Obviously at fixed n_2/n_1 (the dispersions of n_1 , n_2 have negligible importance in the present examples) LD/A is almost independent of A , and LD/A varies almost linearly with n_2/n_1 .

Though not considered when averaging the fields in eq. (1), it is evident (at least with distances of the order of the wavelength) that the magnitude of the effects can be increased if the centre of gravity of the chromophore concentration is moved nearer the dielectric discontinuity surface, in agreement with the larger and more varying WLD of the anions. A complete

Table 1 Wiener linear dichroism of various chromophoric ions in CTAB–water solutions (20% CTAB) with almost complete micellar order, $d=0.1$ cm

Chromophore	Symmetry	Wavelength (nm) of LD _{max} and A _{max}	LD/A
Co(NH ₃) ₆ ³⁺	O _h	340	-0.015
		470	-0.020
Co(en) ₃ ³⁺	D ₃	340	—
		469	—
Fe(dipy) ₃ ²⁺	D ₃	297	-0.013
		530	-0.011
Fe(gmi) ₃ ²⁺	D ₃	552	-0.010
Co(NO ₂) ₆ ³⁻	O _h	258	+0.020
		358	+0.025
Fe(CN) ₆ ³⁻	O _h	428	+0.036
MnO ₄ ²⁻	T _d	310	+0.066
		525	+0.044
CrO ₄ ²⁻	T _d	270	+0.027
		365	+0.020

Concentrations varied between $A=0.1$ – 1.5 (dipy=dipyridyl, gmi=glyoxal-bis-*N*-methylimine, en=ethylenediamine).

theory should thus include some concentration profile function. Effects caused by the large electrostatic fields (C, D in Fig. 2) should also be considered: orientation dichroism by permanent or induced dipoles on the chromophores and anisotropically perturbed transition probabilities due to polarisability¹⁴ of the electronic states participating in the absorption. It is important finally to investigate to which extent the WLD fields are consistent with different local field models^{15–18}.

BENGT NORDÉN
ÅKE DAVIDSSON
L. B. Å. JOHANSSON

Inorganic Chemistry 1,
University of Lund,
Chemical Center,
POB 740 S-220 07 Lund 7, Sweden

Received February 9; accepted April 20, 1976.

- Fraser, R. D. B., *J. chem. Phys.*, **21**, 1511–1515 (1953).
- Sackman, E., and Möhwald, H., *J. chem. Phys.*, **58**, 5407–5416 (1973).
- Davidsson, Å., and Nördén, B., *Chem. Scr.*, **9**, 49–53 (1976).
- Flory, P. J., and Abe, Y., *Macromolecules*, **2**, 335–342 (1969).
- Noda, I., and Hearst, J. E., *J. chem. Phys.*, **54**, 2342–2354 (1971).
- Michl, J., Thulstrup, E. W., and Eggers, J. H., *J. phys. Chem.*, **74**, 3868–3878 (1970).
- Gö, N., *J. phys. Soc. Japan*, **23**, 88 (1967).
- Wiener, O., *Abh. sachs. Akad. Wiss.*, **32**, 509–604 (1912).
- Bragg, W. L., and Pippard, A. B., *Acta Crystallogr.*, **6**, 865–867 (1953).
- Mayfield, J. E., and Bendel, I. J., *Biopolymers*, **9**, 655–675 (1970).
- Copic, M., *J. chem. Phys.*, **26**, 1382–1390 (1957).
- Perutz, M. F., *Acta Crystallogr.*, **6**, 859–865 (1953).
- Johansson, B. Å., Lindblom, G., and Nördén, B., *Chem. Phys. Lett.*, **39**, 128–133 (1976).
- Liptay, W., Weisenberger, H., Tiemann, F., Eberlin, W., and Konopka, G., *Z. Naturf.*, **23**, 377–393 (1968).
- Neugebauer, H. E., *Can. J. Phys.*, **32**, 1–8 (1954).
- Vuks, M. F., *Optics Spectrosc.*, **20**, 361–364 (1966).
- Rohleder, J. W., and Luty, T., *Molec. Crystalllogr.*, **5**, 145–163 (1968).
- Bakhshiev, N. G., Girin, O. P., and Libov, V. S., *Optics Spectrosc.*, **14**, 395–397 (1963).

Dye sensitised zinc oxide: aqueous electrolyte: platinum photocell

TITANIUM oxide photocells, which are photochemically stable and decompose water into hydrogen and oxygen^{1,2}, have a low efficiency³ and their spectral response is mostly in the ultraviolet, where the Sun's output is low. Other semiconductors such as cadmium sulphide or gallium phosphide, having smaller band gaps, are destroyed by light and their photocurrents deteriorate rather rapidly⁴. Photocurrents in dye sensitised photocells composed of aqueous electrolyte solutions and various semiconductor electrodes have also been studied^{5–7}. All results obtained so far show that the dye sensitised photocurrents are very small, $\sim 10^{-6}$ – 10^{-9} A cm⁻², corresponding to a quantum efficiency $\leq 10^{-2}$. In this paper we report the successful construction of cells with a high power efficiency.

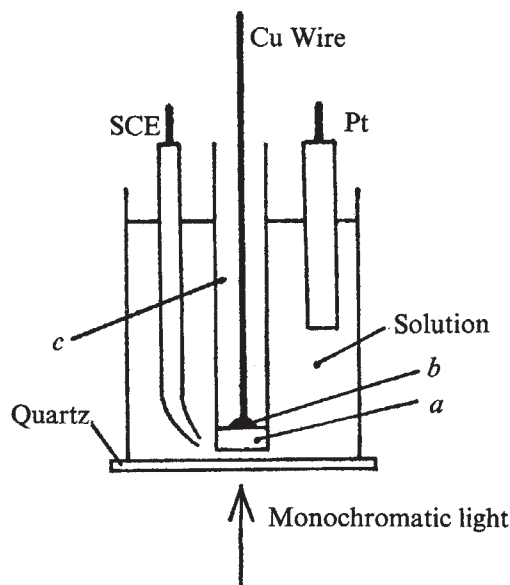
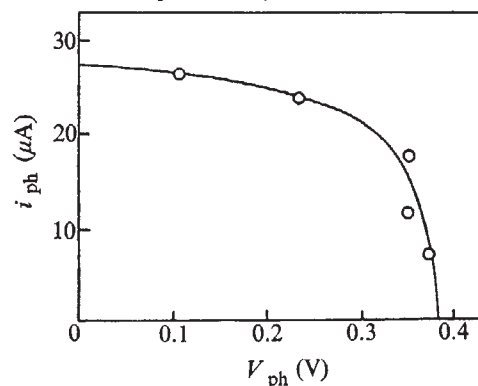


Fig. 1 The design of the photocell. *a*, ZnO sinter disk; *b*, indium contact; *c*, glass tube cemented to the electrode with epoxy resin. SCE, standard calomel electrode.

Most of the work on dye sensitised photocells by previous workers as well as in the early stage of our work was done with the electrodes dipped into aqueous electrolyte–dye solutions. Various experimental results obtained seem to show, as Gerischer and Memming pointed out^{5,8}, that only the dye molecules (or ions) directly adsorbed on the surface of the electrode cause the photocurrent, not those in the bulk of the solution. The poor photocurrents can then be attributed primarily to a small quantity of such a dye on the surface of the electrode. For instance, if the dye forms a monomolecular layer, the absorptivity by the dye is $\sim 10^{-2}$, and therefore most of the light passes through the dye layer without causing any useful effect.

We first tried to coat various insoluble dye films on the surfaces of the electrodes to increase the 'useful' absorbance, but this was unsuccessful, mostly because these films form electrical insulating layers, when they become sufficiently thick to appreciably absorb light. After these trials we succeeded in constructing a cell showing a high power efficiency. The electrode is a zinc oxide sinter disk ~ 1 cm in diameter, prepared from commercial high grade ZnO powder moulded by compression and heated at 1,300 °C for 1 h in air. The disk is then dyed by keeping it in a concentrated aqueous rose bengal solution. We found that the porosity and the dye-adsorbing capability of the ZnO sinter depend largely on the ZnO powder used and the method of sintering. Some adsorb the dye evenly and strongly and others do not. With the former, a high photocurrent efficiency could be obtained. The counterelectrode is a

Fig. 2 Curve of photocurrent (i_{ph}) against photovoltage (V_{ph}) for the ZnO photocell dyed with rose bengal.



plane platinum plate. The solution comprises 0.2 M Na₂SO₄, 0.13 M potassium iodide and 1.0×10^{-3} M iodine.

The cell design is shown in Fig. 1.

Figure 2 shows the photocurrent obtained across a variable resistance, R . The photocurrent on short circuit was 28 μ A which corresponds to a quantum efficiency of 15% (the number of electrons passed per s divided by the number of incident photons per s). The open-circuit photovoltage was ~ 0.4 V, and the maximum power (i^2R) was 6.2×10^{-6} W at $R=20$ k Ω , with 4.1×10^{-4} W, of incident radiation, monochromatised at 563 nm, corresponding to a power conversion efficiency (power produced per s divided by incident light energy input per s) of 1.5%.

These values are at least one order of magnitude higher than our previous values (ref. 8, and H. Tsubomura *et al.*, unpublished), the photocurrent being also quite high compared with those reported before by other authors.

The high efficiency of our cell is probably due to two things. First, the zinc oxide sinter has a fairly low inner resistance, ~ 100 Ω cm. Moreover, as it is porous, the overall quantity of the dye adsorbed on the surface of the electrode can be fairly large, with the thickness of the dye layer on the convoluted surface being kept rather small compared with those obtained by the coating method mentioned before (maybe ~ 1 –2 molecular layers). Thus, we get a large total absorption of light without producing undesirable insulating or energy dissipating effects. Second, the use of a KI–I₂ redox couple improves the cell in two ways. First, the iodide ion serves as an electron supplier to the photo-oxidised dye. Without I[–], or any other reducing agent, the photocurrent decays quickly due to an accumulation of the oxidised dye on the surface⁸. Second, the KI–I₂ couple (or, more correctly, I[–]–I₃[–] couple) keeps the electrode potential of the counterelectrode low and the photovoltage nearly the same as under open-circuit conditions. In other words, the presence of I₂ in the solution prevents the accumulation of over-potential of the platinum electrode at the working condition. (It was confirmed that the dark current of this cell was negligibly small).

The photogenerated power of the cell decreased drastically as the I₂ concentration was lowered. At $[I_2] \leq 7 \times 10^{-6}$ mol l^{–1} the potential of the counter electrode shifts to the negative side, the size of the shift being the larger the higher the photocurrent. We studied the effect of adding other redox couples. With hydroquinone–quinone, for example, a similar i – V curve was obtained, but the I[–]–I₂ couple has given the best result so far.

The merits of our present cell are: (1) The electrode material is zinc oxide sinter which is much less expensive than single crystal semiconductors. It is also very easily prepared in any large size. (2) The dye absorbs light in the spectral region where the solar energy is most abundant. (3) Further improvement of the characteristics of such cells seems promising—by studying optimum conditions for the method of preparation of the semiconductor and the choice of the redox couple, dyes and their concentrations.

H. TSUBOMURA
M. MATSUMURA
Y. NOMURA
T. AMAMIYA

Department of Chemistry,
Faculty of Engineering Science,
Osaka University,
Toyonaka, Osaka, 560 Japan

Received February 13; accepted April 7, 1976.

¹ Fujishima, A., and Honda, K., *Nature*, **238**, 37 (1972).

² Fujishima, A., Kohayakawa, K., and Honda, K., *Bull. chem. Soc. Japan*, **48**, 1041 (1975).

³ Ohnishi, T., Nakato, Y., and Tsubomura, H., *Ber. Bunsenges. phys. Chem.*, **79**, 523 (1975).

⁴ Gerischer, H., *Electroanal. Chem. Interfac. Electrochem.*, **58**, 263 (1975).

⁵ Gerischer, H., *Photochem. Photobiol.*, **16**, 243 (1972).

⁶ Memming, R., *Photochem. Photobiol.*, **16**, 325 (1972).

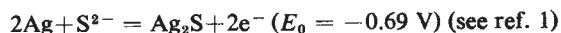
⁷ Fujishima, A., Watanabe, T., Tatsuoki, O., and Honda, K., *Chem. Lett.*, **13**, (1975).

⁸ Matsumura, M., Yamamoto, K., and Tsubomura, H., *J. chem. Soc. Japan*, **399** (1976).

Photoelectrochemical energy conversion and storage using polycrystalline chalcogenide electrodes

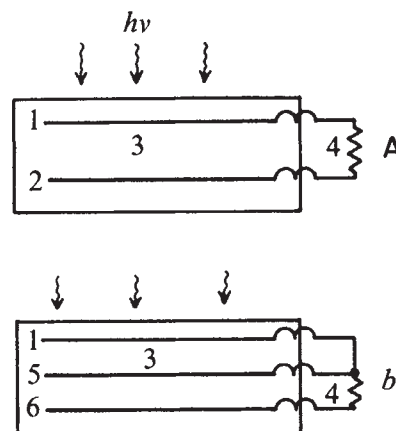
WE report here on a major improvement in the conversion efficiency of corrosion-free photoelectrochemical cells (PECs) and on a novel extension of such cells which allows the storage of part or all of the converted energy *in situ* for subsequent use.

We have used, for example, a 2 cm² Cd–Se photoelectrode, prepared by the electrolytic codeposition of Cd and Se on a conducting base, followed by heat treatment in an inert atmosphere. Corrosion of the photoelectrode was prevented by using an aqueous or organic solution of sulphide in which some elemental sulphur was dissolved and from which air was rigorously excluded. Using a Cd–Se electrode, currents of 7–10 mA cm^{–2}, and open circuits, potentials from 450–560 mV were obtained in AM1 sunlight. Because of the large currents generated with these electrodes, the counterelectrode may limit the current, so that minimally polarised counterelectrodes with large surface areas, such as activated carbon, should be used. Part of or all the converted energy is stored in a controlled way in the system by the introduction of an electrode of porous silver:



All known stable PECs use semiconductors with a wide bandgap ($E_g > 2.8$ eV) as the photoactive electrode^{2–4}, thus severely limiting their practical use for solar energy conversion. The main difficulty in the use of semiconductors with bandgaps which are optimal for solar energy conversion ($E_g \sim 1.4$ eV) in PECs, is corrosion of the photoactive electrode, either directly or under the influence of light (unpublished). The aim of most workers in this field has been the photoelectrolysis of water^{5–7} which we do not consider a practical proposition (see ref. 8). For the photoelectrolytic production of hydrogen from water only the O/O^{2–} redox couple can be used. If, however, the purpose is electricity generation, other redox couples such as S/S^{2–}, with a much lower redox potential, can be used (E_0 (O/O^{2–}) = 0.4 V; E_0 (S/S^{2–}) = –0.45 V; both in 1N basic solution). Therefore, the photoelectrode can be held at a much lower anodic potential, aiding in the prevention of corrosion. Besides Cd–Se, other semiconductors such as Cd–S,

Fig. 1 PEC and PESC configurations: a, 2-electrode system (current direction in the light opposite to that in the dark, when used as storage cell); b, 3-electrode system (no current reversal between light and dark). 1, Chalcogenide photoelectrode; 2, inert counterelectrode or storage electrode; 3, S^{2–}/S solution; 4, load; 5, storage electrode; 6, inert counterelectrode.



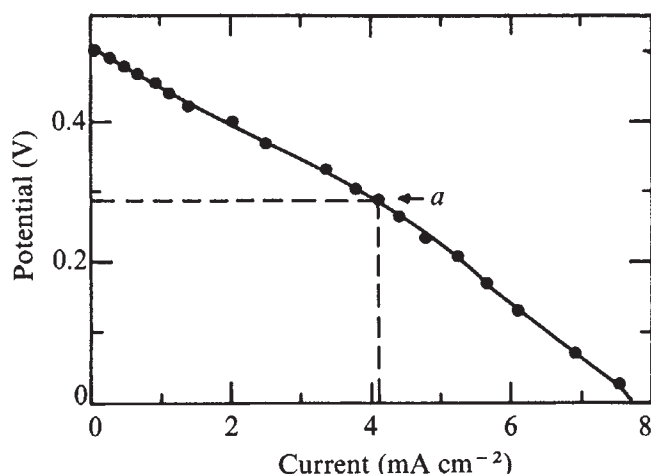


Fig. 2 Voltage-current behaviour of a PEC comprising a 1 cm² polycrystalline Cd-Se layer on Ti and a 4 cm² counter-electrode of activated carbon, both immersed in a 1M solution in OH⁻, S²⁻ and S. Illumination by a 1A tungsten-iodine projector equivalent to AM1 sunlight. a, Maximum power.

Cd-Te (n- and p-type), Zn-Se or Bi₂S₃ are also stable as photoelectrodes in a polycrystalline form in such conditions.

Figure 1 shows two configurations of a chalcogenide PEC, or photoelectrochemical storage cell (PESC). Such cells are stable over periods of months (longer term trials are in progress). The absence of corrosion has been demonstrated in two ways. A 1 cm² polycrystalline Cd-Se photoelectrode on Ti was prepared by the electrolytic codeposition of Cd and Se from an acidic aqueous solution of CdSO₄ and SeO₂ using 4.5 C per electrode side, yielding a Cd-Se layer of about 7.5 μmol. The electrode was immersed in 10 ml of 1M S²⁻/S solution, and illuminated for 15 h. Enough current (270 C) was passed to corrode all of the Cd-Se layer about 200 times, but the performance of the cell did not diminish. Moreover, the solution in the cell was analysed for Cd, before and after illumination, using atomic absorption spectroscopy. No change in the Cd content of the solution (< 2 × 10⁻⁶ M) was observed. During the first few hours of illumination, a thin layer of the chalcogenide (for example, Cd-Se) will undergo partial S/Se substitution if non-sulphide chalcogenides are used. Thus, a thin sulphide layer is obtained near the surface exposed to the electrolyte. Electron microprobe studies show that the S/Se ratio, in a Cd-Se electrode that had been illuminated for several hours, was about 1:1 in the top micrometre of the layer. Also, X-ray diffraction

showed the presence of some CdS in the Cd-Se layer. The spectral response of the electrode remains essentially that of the original chalcogenide, however, indicating that only a top layer of Cd-Se was transformed into CdS.

The performance of a Cd-Se PEC is illustrated in Fig. 2. It can be seen that the potential drops rather severely under load. This is attributable to resistance losses in the system such as those in the photoactive layer, those in the electrolyte (which depend on the cell geometry, among other things), polarisation losses at the counterelectrode, and losses at the semiconductor-metal contact. The necessary improvement in the conductivity remains as a major challenge in obtaining a practical PEC.

Results obtained using a 3-electrode PESC are shown in Fig. 3. Storage efficiencies are low when the storage electrode is in the same solution as the other electrodes. If a cation-specific membrane is used to separate them, and care is taken to exclude sulphur from the storage compartment, storage current efficiencies of up to 90% can be obtained with a silver electrode. It is possible to use other systems, such as FeS-Fe₂S₃ and Sn-SnS. By using Fe(OH)₂-FeO(OH) as a storage electrode a PESC with an oxide electrode and O/O²⁻ couple can be constructed.

A report to be published elsewhere will consider the chemical and physical basis of the stability and storage capabilities of these cells, their spectral response, potentiostatic behaviour and the analogy between the semiconductor-electrolyte boundary and Schottky and Bardeen semiconductor-metal contacts, particularly with regard to surface states.

GARY HODES
JOOST MANASSEN

Department of Plastics Research,

DAVID CAHEN

Department of Structural Chemistry,

Weizmann Institute of Science, Rehovot, Israel

Received March 15; accepted April 28, 1976.

¹ Latimer, W. M., *The Oxidation States of the Elements and their Potentials in Aqueous Solutions*, 2nd ed. (Prentice-Hall, New York, 1952).

² Hodes, G., Cahen, D., and Manassen, J., *Nature*, (in the press).

³ Fujishima, A., and Honda, D., *Nature*, **238**, 37-38 (1972).

⁴ Wrighton, M. W., Morse, D. L., Ellis, A. B., Ginley, D. S., and Abrahamson, H. B., *J. Am. chem. Soc.*, **98**, 44-48 (1976).

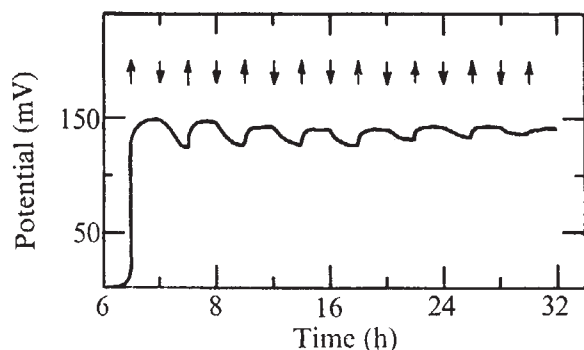
⁵ Gerischer, H., *J. electroanal. interf. Electrochem.*, **58**, 265-274.

⁶ Fujishima, A., Kohayakawa, K., and Honda, K., *J. electrochem. Soc.*, **122**, 1487-1489 (1975).

⁷ Mavroides, J. G., Tchernev, D. I., Kafalas, J. A., and Kolesar, D. F., *Mater. Res. Bull.*, **10**, 1023-1030 (1975).

⁸ Yoneyama, H., Sakamoto, H., and Tamura, H., *Electrochim. Acta*, **20**, 341-345 (1975).

Fig. 3 Voltage against time for a 3-electrode PESC without membrane. The PESC consists of the components of the cell of Fig. 2 to which an Ag-Ag₂S storage electrode is added in the configuration illustrated in Fig. 1b. Load, 680 Ω; ↑, light on; ↓, light off. Before the experiment the cell was completely discharged. The overpotential for sulphur reduction at the photoelectrode is sufficiently high to ensure that almost all of the current flows from the storage electrode over the load to the counterelectrode in the dark and not back to the photoelectrode.



Oldest organic remains of boring algae from Polish Upper Silurian

REMNANTS of cells and branched filaments of carbonate boring (endolithic) algae have been found in the Upper Silurian sedimentary rocks of eastern Poland. The material was obtained from a borehole at Widowo, near Bielsk Podlaski at a core depth of 524-525 m. The age of the sediments at a core depth of 545-585 m was determined as the Upper Silurian (Wenlockian), based on tabulate corals and stromatoporoids. The coral-stromatopore assemblage corresponds to the Jaagarahu Stage of the Silurian of Estonia^{1,2}. It includes tabulate and heliolitoid corals *Coenites juniperinus* Eichwald, *Palaeofavosites collatatus* Sokolov, *P. frivulus* Klaaman, *P. tersus* Klaaman and *Heliolites decipiens* M'Coy³, and stromatoporoids *Densastroma pexisum* Yavorsky and *Stromatopora impexa* Nestor (Kazmierczak, unpublished).

The boring algae occur in poorly sorted, coarse calcarenite, comprising mostly fragmented crinoid columnals and finely broken brachiopod and mollusc coquina. Pieces of the dasycladaceous alga *Rhabdoporella* Stolley, and bundles of *Girvanella* Nicholson and Etheridge are common. A few fragments of trepostome bryozoans and solenoporeacean algae have also been observed.

Carbonate boring algae have been found exclusively within the crinoid columnals. Since they penetrate the columnals from all directions including the lumen and suture sides, it is concluded that their colonisation took place after the disintegration of the crinoid stems, but before their burial and the subsequent consolidation of the sediments.

The algae were studied *in situ*, in petrographic thin sections (Fig. 1a). The borings begin at the surface and penetrate perpendicularly into the interior of crinoid columnal fragments. No epilithic filaments have been observed. The endolithic filaments are composed of torulose, isodiametric, or slightly elongated cells 10–30 μm wide (mean $\pm$ s.d. = 18.1 ± 4.6 , $n=100$), and 12–35 (22.2 ± 6.0 , $n=100$) μm long. They are repeatedly and densely branched, forming small bushes, 80–200 μm in diameter (105.6 ± 35.0 , $n=20$), within the substrate. Branching in all directions is lateral or sub-dichotomous. The cellular outlines are well defined by a brown, slightly shrivelled wall structure. One or two

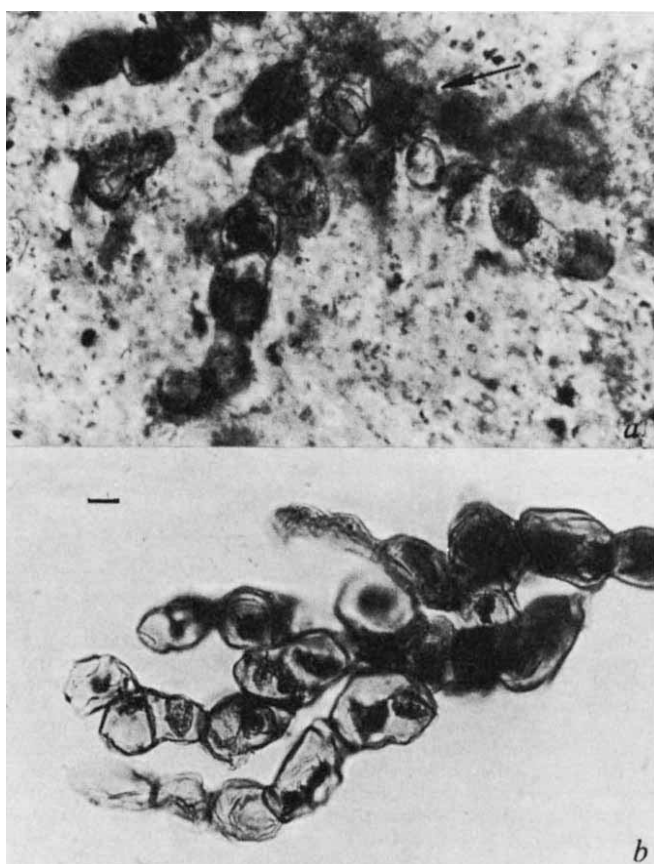


Fig. 1 Branched filaments of the Upper Silurian boring green alga discussed here: *a*, boring filaments *in situ*, penetrating a columnal crinoid fragment, petrographic thin section, transmitted light (arrow, substrate surface); *b*, filaments extracted with dilute HCl. Scale bar, 10 μm .

irregularly shrunken or flattened bodies $4\text{--}10 \times 6\text{--}14 \mu\text{m}$ ($6.3 \pm 2.5 \times 8.1 \pm 4.3$, $n=20$) are found within most cells.

Algal filaments have been extracted by gradual dissolution of the carbonate matrix with 0.1 N HCl, and mounted in glycerol. They remained cohesive, although rigid and fragile (Fig. 1b). Selected filaments have been embedded in Araldite and sectioned for transmission electron microscopy (TEM), with and without osmium tetroxide, lead citrate and uranyl acetate treatment. Application of osmium tetroxide did not affect the image, but lead citrate and uranyl acetate increased the contrast.

At the TEM level, the outer walls, 300–1,200 nm thick, revealed a multilayered structure with distinguishable,

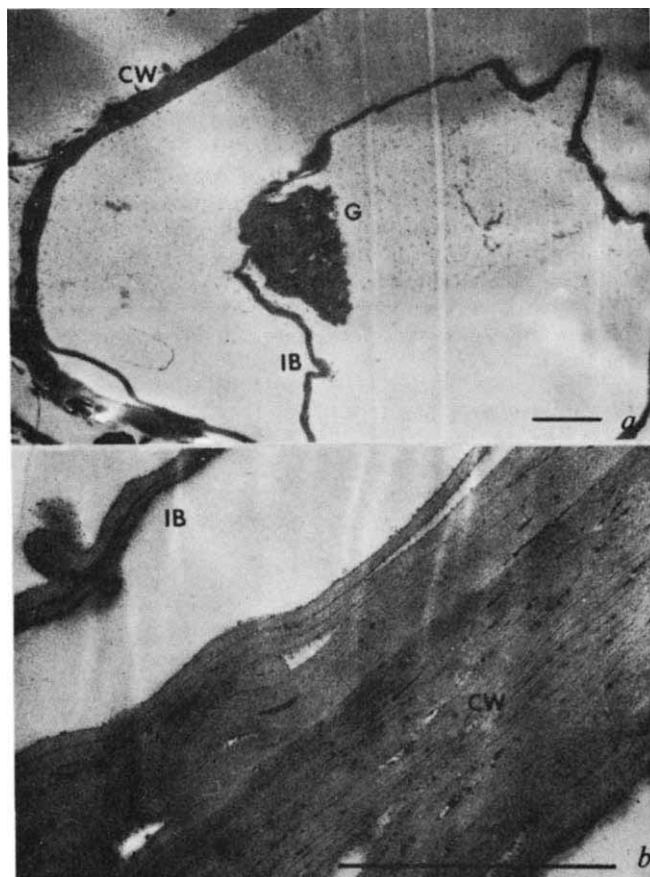


Fig. 2 Preserved cellular structures of the Upper Silurian boring green alga: *a*, section through the cell ($\times 9,200$); *b*, detail, multilayered wall structures (CW, outer cell wall; IB, internal body outline; G, granular aggregate), ($\times 36,800$). Scale bars, 1 μm .

individual, layers 10 nm thick (Fig. 2a and b). Packets of layers are frequently folded or detached. Cross walls have similar structures, but are generally thinner (100–400 nm). They meet with the longitudinal walls by forming a triangular space. No plasmodesmic connections have been identified on the cross walls. The outlines of the internal bodies are less than 100 nm thick. They also have a layered substructure and an electron density similar to the external and cross walls, but they are more convoluted and shrivelled. Occasional thickenings extend towards the lumen of the internal bodies. Within the internal bodies there are occasional isodiametric or flattened aggregations of electron dense masses with granular substructures. The ultrastructure of the preserved organic residue permits only an identification of the multilayered cell walls. Subtle structural features such as plasmodesmata, considered an important taxonomic marker⁷, are either not present or have become diagenetically obliterated. No closer identification of intracellular features has been possible.

Borings of endolithic algae are known as far back as the Ordovician⁴. The significance of the present finding lies in the preserved organic parts which allow determinations of cell sizes and shapes, the positions of cross walls and modes of branching, and the fine structures of the organic residues. The preserved features, along with comparisons with modern boring algae, identify the remains as Chaetophorales.

Endolithic algae are common within the illuminated portions of modern marine carbonate rocks and skeletal fragments in sediments⁸. Extant endoliths are represented by a few highly specialised prokaryotic (cyanophytes) and eukaryotic (chlorophytes and rhodophytes) algae. The

Silurian fossil described here compares well with septate endolithic chaetophoralean chlorophytes such as *Phaeophila*, *Eugomontia* and *Entocladia*⁶, but cannot be directly identified with any of the modern genera. Morphologically, it is closest to *Entocladia* (or some growth forms of *Eugomontia*), which does not bear bristles (chaetae). The torulose cell shape and the mode of branching are similar to various epilithic (*Gongrosira*) as well as endolithic chaetophorales. Thus, it is conceivable that this Silurian endolith represents an ancestral type of some Holocene endolithic green algae.

This work was supported by the Polish Academy of Sciences, the Alexander von Humboldt Foundation (FRG), and the US NSF.

JÓZEF KAZMIERCZAK

Polska Akademia Nauk,
Zakład Paleozoologii,
Al. Zwirki i Wigury 93,
02-089 Warsaw, Poland

STJEPKO GOLUBIĆ

Department of Biology,
Boston University,
2 Cummington Street,
Boston, Massachusetts 02215

Received March 1; accepted April 15, 1976.

- ¹ Nestor, H. *Wenlockian and Ludlovian Stromatoporoidea of Estonia*, ("Valgus", Tallinn, Russian with English summary, 1966).
- ² Kilaaman, E., in *The Silurian of Estonia* (edit. by Kallio, D. L.), 114–125 ("Valgus", Tallinn, Russian with English summary, 1970).
- ³ Stasinka, A., *Acta palaeont. pol.*, 19, 501–517 (1974).
- ⁴ Hessland, I., *Bull. geol. Instn Univ. Uppsala*, 33, 409–928 (1949).
- ⁵ Golubic, S., Perkins, R. D., and Lukas, K. J., in *Study of Trace Fossils* (edit. by Frey, R.), 229–259 (Springer, New York, 1975).
- ⁶ Nielsen, R., *Bor. Tidsskr.*, 67, 245–269 (1972).
- ⁷ Stewart, K. D., Mattox, K. R., and Floyd, G. L., *J. Phycol.*, 9, 128–141 (1973).

Fossil iron bacteria may be preserved in Precambrian ferroan carbonate

CONCLUSIVE evidence for microbial participation in the genesis of the massive sedimentary iron formations from the Precambrian is fragmentary. Structures resembling fossilised bacteria have been reported previously, and they tend to resemble modern iron bacteria such as *Sphaerotilus*, *Gallionella* and *Metallogenium*^{1,2}. The atmosphere at that time is believed to have been almost devoid of oxygen³. The concentration of atmospheric oxygen probably ranged from 0.0001 to 0.001 % of the present atmospheric level, and may have remained at this level for 2×10^9 yr after the emergence of algal photosynthesis⁴. The fossils described in this paper were found during a study of the mineralogy of the Lower Proterozoic Brockman Iron Formation (Hamersley Group) of Western Australia (about 2×10^9 yr old⁵). Two specimens (PM9 and PM24) collected from the Whaleback Shale Member, at Wittenoom Gorge, were identified as dolomitic chert by X-ray diffraction. Optical microscopy showed the presence of fibrous stilpnomelane, and the staining technique suggested by Dickson⁶ distinguished the dolomite as ferroan.

During our study of the carbonates, it was observed that some ferroan dolomite crystals showed filamentous structures on or near the surface of the crystals (Fig. 1b–e). They were stained deep blue, showing the presence of ferrous iron. The morphology of these filaments resembles the modern iron bacterium *Sphaerotilus*.

A weakly acid solution of potassium ferricyanide $K_3Fe(CN)_6$ was used to stain petrographic thin sections (0.2 g $K_3Fe(CN)_6$ in 100 ml 1.5 % HCl)⁶. $K_3Fe(CN)_6$ reacts with ferrous iron to give a deep blue—Turnbull's blue—precipitate of ferrous ferricyanide:



The section was first cleaned with 'Teepol' to remove all grease or immersion oil, dipped in a beaker containing the above stain,

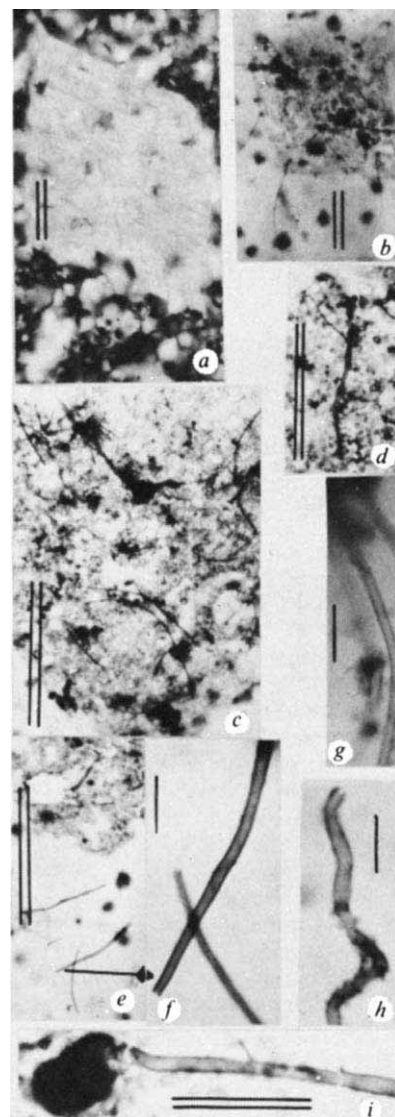


Fig. 1 Except where otherwise stated, all the illustrations are photomicrographs taken on a Zeiss photomicroscope in the Geology Department, Royal School of Mines, London, from thin sections of Whaleback Shale sample PM24. a, Dolomite rhomb before staining under crossed nicols. b, Dolomite rhomb after staining showing filaments on the surface of the crystal. c and d, Filaments on the surface of the crystal, some attached to rounded bodies. e, Filament detached from the crystal during coverslip mounting. Arrow indicates same filament in f at higher magnification. g, Modern specimens of *Sphaerotilus* from Silwood Lake, Berkshire. h, Isolated filament for comparison with g. i, Rounded body to which the filament has been attached, mainly composed of ferrous iron. Single marker bar indicates 10 μm. A pair of marker bars indicates 100 μm.

and allowed to react for 45–60 s. The slide was removed, and immediately rinsed with distilled water. The stain is on the surface of the minerals in the form of a precipitate, and may easily be destroyed if touched. Thorough washing with water is necessary to remove all the acid; otherwise carbonate crystals (especially calcite) would dissolve. Care is also required in the use of water, as prolonged contact of the precipitate with water renders it soluble.

Isolated filaments (f and h) were obtained when mounting the coverslip. A drop of glycerine was applied to a thoroughly water-washed, stained slide, to preserve it for a longer period of time. The slide was then warmed to make the glycerine less viscous, and the coverslip put in place. During this operation, some of the filaments became detached from the substrate and moved away to a new position free of the matrix.

Almost all the filaments I observed were attached to or lying on the surface of ferroan dolomite crystals, and each filament was filled with ferroan carbonate. No filaments were found in the rest of the matrix of the rock. The length of the filaments, on average, was 100 μm , and the fossils show no apparent morphological distortion, except that some of them had been broken during coverslip mounting. Many of the filaments seemed to be attached to a round body mainly composed of ferroous iron, to judge from the intensity of the staining (Fig. 1i).

Before staining (for details, see Fig. 1), the filaments in the slide were invisible (Fig. 1a). This may mean that they were preserved as carbonate moulds formed during early diagenesis. This is an example of an authigenic preservation process⁷, involving early cementation in soft sediment, usually by iron and carbonate compounds—in this case, ferroan dolomite—preserving the configuration of organic parts. The process results in the preservation of structures of delicate shape and form.

The fossil material was compared morphologically with modern species of *Sphaerotilus*, collected from water samples from Silwood Park Lake, Ascot, Berkshire, UK. A few drops of stain were allowed to react with the modern sample, and a similar colour change was observed (Fig. 1g). The morphological agreement is self evident when Fig. 1g (modern bacteria) and Fig. 1f and h (isolated fossil filaments) are compared. Somewhat similar structures have been detected preserved in haematite and chert of Tertiary age⁸, and from other Precambrian iron formations from Canada⁹, and from Western Australia¹⁰. It has been suggested that these other structures^{8–10} represent the remains of fossil iron bacteria. Because of the strong morphological similarity between the stained ferroan dolomite filaments and modern *Sphaerotilus*, I suggest that the structures discussed here also represent the remains of Precambrian iron bacteria.

It might be argued that these tubular structures have formed as the result of boring by algae and fungi, but one would expect boring algae and fungi to be generally distributed throughout the rock. This may be ruled out, however, in the present case, because the filaments are invariably associated only with the ferroan dolomite crystals, and are not generally distributed. Furthermore, algal and fungal filaments are generally much larger, and may often be septate. Septa were not observed in my material. If the structures represent borings, then a trail of mucilaginous material might be expected around the filaments. None were observed, and no trails or 'borings' (if such they are) seem to connect up the individual and spatially isolated ferroan dolomite crystals. Note also that if the filaments do represent the remains of iron bacteria, that present-day examples are always associated with clots of colloidal iron hydroxides; it may be suggested, therefore, that the ferroan dolomite crystals represent the diagenetic products of such clots.

Statements have been made in the past that "bacteria do not have hard parts that can be fossilised, so there is little to be learned directly about their evolution"¹¹. In view of the application of this staining technique to carbonate rocks, this statement must be modified. Also the technique, or similar methods, must be extended to other forms of preservation—such as in pyrite, in limonite, in phosphate, in silica, and in calcite. As well as providing evidence for the presence of iron bacteria in ancient rocks, the technique offers the possibility of discovering previously undetectable microfossils from many different sedimentary environments in the Precambrian that are at present considered to be barren.

I thank Dr M. D. Muir and Mr C. J. Peat for discussions, and Messrs K. A. Plumb, J. Baldwin and R. Giddens for assistance.

SURESH N. KARKHANIS

Department of Geology,
The Royal School of Mines,
Prince Consort Road,
London SW7 2BP, UK

Received February 16; accepted April 27, 1976.

- ¹ Cloud, P. E., Jr, *Science*, **148**, 27–35 (1965).
- ² Tyler, S. A., and Barghoorn, E. S., *Science*, **119**, 606–608 (1954).
- ³ Cloud, P. E., Jr, *Evolution and Environment* (edit. by Drake, T.), 72 (Yale University Press, 1968).
- ⁴ Berkner, L. V., and Marshall, L. C., *Proc. natn. Acad. Sci. U.S.A.*, **33**, 1215–1225 (1965).
- ⁵ Compston, W., and Arriens, P. A., *Can. J. Earth Sci.*, **5**, 561–583 (1968).
- ⁶ Dickson, J. A. D., *Nature*, **205**, 587 (1965).
- ⁷ Schopf, J. M., *Rev. Palaeobotan. Palynol.*, **20**, 27–53 (1975).
- ⁸ Muir, M. D., Hamilton, L. H., Grant, P. R., and Spicer, R. A., in *Microprobe Analysis as Applied to Cells and Tissues* (edit. by Hall, T., Echlin, P., and Kaufmann, R.), 33–58 (Academic, London and New York, 1974).
- ⁹ Barghoorn, E. S., and Tyler, S. A., *Science*, **147**, 563–575 (1965).
- ¹⁰ Walter, M. R., *Abstr. First Australian Geol. Convention, Proterozoic Geology*, Adelaide, 15–16 (1975).
- ¹¹ Broed, R. S., in *Bergey's Manual of Determinative Bacteriology*, 4–14 (Williams and Wilkins, Baltimore, 1975).

Viable bacterial spores recovered from an archaeological excavation

THERMOPHILIC actinomycetes grow rapidly and sporulate profusely in natural high temperature habitats such as composts, stored fodders and cereals. In these habitats can be found large numbers of *Thermoactinomyces* spp. endospores which exhibit longevity and extreme resistance to adverse conditions. Such spores have been recovered from lake muds and have been estimated to be more than 1,000 yr old, but it is important to have more accurate dating. For this purpose we have examined occupational debris from a Roman archaeological site which has been shown to be rich in plant debris and has been accurately dated from associated artefacts. Our findings show a strong correlation between viable spore count of *Thermoactinomyces* spp. and the nature of the substrate. We have also shown that it may be possible to undertake qualitative and quantitative analyses, using microbial taxa with specific environmental requirements, for archaeological interpretation. Furthermore, the study of evolutionary trends within the genus *Thermoactinomyces* is now possible.

Pre-Hadrianic buildings (85–125 AD) outside the conjectured area of the Roman fort and *vicus* at Vindolanda, Northumberland, UK (grid reference 35(NY)/771662; altitude 160–165 m) were located by chance while Mr R. Birley, director of archaeological excavations, was preparing a drain in late 1972. The digging season in 1973 revealed the full potential of this find: within the occupational debris uncovered were two distinct layers which have since been dated as (85–) 90–95 AD and 95–105 (–110) AD = strata I and II respectively in Fig. 1, each of which is characterised by a high organic content (61–89% by dry weight). Two further major occupational levels—strata III and IV in Fig. 1—are superimposed on these layers and can be dated between (105–) 110 and 125 AD; minor debris or occupational levels, for example, stratum IIa in Fig. 1, were encountered occasionally.

The occupational debris levels are sandwiched between compacted clay levels, indicating that the floor of the particular building had degenerated to such an extent that a new and more permanent floor was the answer to such problems as poor sanitation and a rising water table. A new floor may have coincided with partial or total reconstruction and/or replacement of the superstructure, for the basal plans of the successive building programmes provide a complex picture. The walls of the buildings were composed essentially of birch uprights interspersed with seasoned and fashioned oak; the whole was interwoven with birch branches, often supplemented by other tree species (for example, hazel, oak and willow).

The rooms so far excavated, which cover 30 m², were possibly put to various uses during the 40–45 yr of occupation: rooms which probably served in the first instance as living-quarters later had other functions such as bracken storage, tanning and leather-working, animal housing and the disposal of occupational debris. An earth and rubble

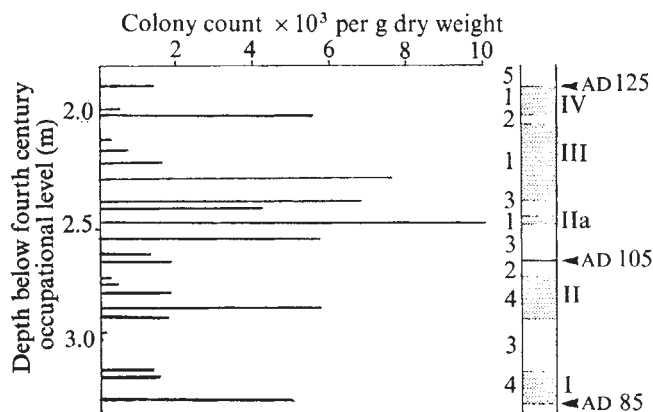


Fig. 1 Recovery of thermoactinomycetes from occupational debris and compacted clay strata at Vindolanda, Northumberland. Diluted samples were spread on the dried surface of a selective isolation agar containing novobiocin and cycloheximide⁷, and incubated at 50 °C for 48 h. I–IV, Major occupational strata. Composition of strata: 1, highly organic, composed mainly of straw and bryophytes; 2, less organic, with high percentage of mineral soil; 3, compacted clay; 4, highly organic, composed mainly of bracken, straw and bryophytes; 5, earth and rubble infill.

infill around 125 AD delimits the end of this phase of occupation; the site was redeveloped from the late second century through to the fourth century.

The clay compactions, between the different phases in the pre-Hadrianic occupational levels, have created pockets of chemical conditions which have preserved a wide range of natural and man-made materials representative of the period. Preservation is due mainly to the anaerobic conditions, probably enhanced by a rising water table, and to chemicals derived from the medium. The latter is highly organic, being composed of bracken¹, straw, bryophytes², puff-balls³, hazel and walnut shells, acorns, gorse pods, leaves, stems (mainly heather) and twigs (mainly ash, hazel, oak, pine, rowan and willow), as well as a wealth of animal material which includes a considerable number of stable-fly puparia, bones, hairs, feathers, oyster and mussel shells, and excreta (part of which is human in origin)⁴. Tannins, derived from many of these components of the medium and from the considerable quantity of leather material present, are anti-bacterial in nature. And tannins may have been produced purposely for leather-work. This is suggested by the extensive number of leather offcuts present and the urine-impregnated medium, which may have resulted from the collection of urine for tanning. Vivianite is common, and the enormous quantity of bones present in the deposits contribute phosphate. pH measurements of all deposits so far examined are remarkable for their homogeneity, all being within the limited range of 5.3–5.6.

As well as leather boots, shoes, sandals, clothing and offcuts, this medium has preserved a unique collection of writing tablets made from thin slivers of wood and bearing carbon-based ink writing, together with considerable quantities of cloth.

An examination of the levels of occupational debris for artefacts and macro-subfossils and other environmental material is being supported by microbiological analyses. Although palynological investigations are just beginning, the recovery of viable bacterial endospores from the complex media described above is considered here.

The pre-Hadrianic occupational layers seemed to be an ideal substrate in which to look for *Thermoactinomyces* endospores since those of *T. vulgaris* have been isolated from present-day bracken litter⁵, are commonly associated with stored cereals and straw⁶ and exhibit longevity⁷.

Sections about 13 cm long (about 700 cm³ in volume)

were cut from a freshly exposed vertical face of an excavation trench wall. These were stored in tightly sealed polythene jars at room temperature. Samples for microbial analysis were removed from the centre of each section after slicing it from top to bottom with a sterile knife. High numbers of *Thermoactinomyces* spp. were recovered from the occupational layers and smaller numbers from the compacted clay between these layers and from the infill above the deposits (Fig. 1). Since stratum I can be dated between 85 and 95 AD, there is strong evidence that spores have survived 1,890 yr. The chemical conditions of the deposits have contributed to the survival of viable *Thermoactinomyces* endospores. The persistence of biodegradable materials in these deposits suggests low or non-existent microbial activity brought about by anaerobic conditions, a high water table and a low temperature. There is a highly significant correlation between the colony count and the presence or absence of occupational strata (Fig. 1).

It seems therefore that counts of microbial taxa with specific environmental demands may prove valuable in archaeological interpretation. In the case described above, the high temperature and oxygen requirements of *Thermoactinomyces* spp. reflect the prevailing environmental conditions within the organic media before the suppression of endospore germination by encapsulation of the deposits.

Thermoactinomyces belonging to the species *T. vulgaris*, *T. sacchari*, *T. dichotomica* and *T. candidus* have been identified among the isolates⁸. Taxonomic investigations have also revealed that the relative frequency of the *Thermoactinomyces* spp. differed markedly between the pre-Hadrianic deposits and recent soil strata. Furthermore, strains of both *T. vulgaris* and *T. dichotomica* are sometimes found within the deposits which differ from those considered typical of the currently recognised species⁸. The palaeogenetic implications of these findings are being investigated in more detail.

This work was supported by the NERC.

M. R. D. SEAWARD
T. CROSS
BRIDGET A. UNSWORTH

Schools of Environmental Science and
Biological Sciences,
University of Bradford,
Bradford BD7 1DP, UK

Received February 13; accepted April 27, 1976.

- ¹ Seaward, M. R. D., *Bot. J. Linn. Soc., Lond.* (in the press).
- ² Seaward, M. R. D., and Williams, D., *J. Archaeol. Sci.* (in the press).
- ³ Watling, R., and Seaward, M. R. D., *J. Archaeol. Sci.* (in the press).
- ⁴ Seaward, M. R. D., *The Vindolanda Environment* (Barcombe Publications, Halthwistle, 1976).
- ⁵ Cross, T., *J. appl. Bact.*, 31, 36–53 (1968).
- ⁶ Lacey, J., *Ann. appl. Biol.*, 69, 187–212 (1971).
- ⁷ Cross, T., and Attwell, R. W., in *Spore Research 1973* (edit. by Barker, A. N., Gould, G. W., and Wolf, J.), 11–20 (Academic, London and New York, 1974).
- ⁸ Unsworth, B. A., Cross, T., Seaward, M. R. D., and Sims, R. E., *J. appl. Bact.* (in the press).

Accumulation of cadmium by the American oyster, *Crassostrea virginica*

CADMIUM is a major environmental pollutant potentially harmful to health, and if the sea becomes polluted with this metal there could be a reduction in extensive sources of food¹. Seafoods constitute a source of cadmium in the human diet² and in view of abundant evidence that shellfish accumulate trace metals^{3–6}, it is important to investigate cadmium pollution. People have become ill from cadmium poisoning after ingesting foods containing concentrations of 13–15 µg g⁻¹ (13–15 p.p.m., ref. 7). We now report that adult oysters reared in seawater containing 0.005 p.p.m. cadmium accumulated up to 10.75 p.p.m. in 40 weeks. This accumulation, plus cadmium

Table 1 Cadmium in whole meats of *Crassostrea virginica* after exposure to 5 parts per 10⁶ Cd²⁺ for 40 weeks (November 1972 to August 1973) in flowing unfiltered seawater

Date	Time (weeks)	Mean biweekly temperature (°C)	µg Cd per g wet weight		µg Cd per g dry weight	
			Control	Experimental	Control	Experimental
1973						
November	0	10.0	2.72	2.72	11.83	11.83
December	4	4.3	2.48	3.60	13.10	21.75
1974						
January	8	2.8	2.58	4.02	15.57	22.00
February	12	3.2	3.11	4.14	19.35	24.90
March	16	4.3	2.10	4.96	13.42	28.55
April	20	7.2	2.62	5.29	13.60	27.59
May	24	11.1	2.81	4.76	16.75	33.77
June	28	16.4	1.93	5.71	10.85	39.22
July	32	20.2	1.65	9.10	12.19	63.27
August	36	21.1	1.76	10.37	17.47	85.83
August	40	22.6	1.69	13.57	12.55	104.91

naturally present, brought the concentration of cadmium to 13 p.p.m. in the soft tissue, which represents a potential health hazard if oysters constitute a major item of the diet.

Three-year-old oysters, *Crassostrea virginica*, were obtained from Long Island Sound in October 1972. The mean shell length was 10.2 cm with a mean whole wet meat weight of 8.1 g. The animals were acclimatised for 1 month, in six fibreglass troughs, each measuring 3.75 m × 30 cm × 25 cm deep and supplied with flowing unfiltered seawater (30‰ salinity, ambient temperature of 1.5–23 °C) at a rate of 16 l min⁻¹. Each trough held 50 oysters on a false bottom consisting of polyethylene grids resting on polyvinylchloride pipe about 2.5 cm above the bottom.

After acclimatisation of the oysters, three troughs received a cadmium chloride solution (CdCl₂ · 2.5 H₂O) and three troughs served as controls. The seawater-cadmium stock solution contained 0.016 mg Cd per ml and was pumped at the rate of 5 ml min⁻¹ to give, when mixed with incoming seawater entering at a rate of 16 l min⁻¹, a calculated concentration of 0.005 p.p.m. cadmium. Samples of water were taken from the troughs weekly, concentrated by a Chelex resin technique developed by Dr Earl Davey of our laboratory and assayed by atomic absorption spectroscopy. A range of values from 0.0047 to 0.0053 p.p.m. was found in the seawater from the experimental troughs and 0.0001–0.0002 p.p.m. was found in the control samples. Throughout this study, the animals received no supplementary food because the seawater supplied to the troughs apparently contained sufficient food organisms. Mortality was negligible in both control and experimental groups.

Every 4 weeks, five oysters were removed from the experimental and control groups. They were shucked and the shell liquor and body fluids were retained with the whole wet meats. The meats were weighed, dried at 100 °C for 24 h and ashed at 425 °C for 18 h. After cooling, 10 ml of concentrated HNO₃ was added to the ashed samples which were then dried at 100 °C for 24 h and ashed for 18 h. When necessary, ashing and treatment with HNO₃ were repeated until a clean white ash was obtained. The final ash was dissolved in 5% HNO₃ and analysed for cadmium by atomic absorption spectroscopy. The results were expressed as µg of cadmium per g (p.p.m.) for both wet and dry weights.

A mean of 2.72 µg cadmium per g wet weight was found in oysters immediately after collection. Concentrations of cadmium in seawater at the sampling site were 0.0001–0.0002 p.p.m. Although the dry weight data show that *C. virginica* can accumulate cadmium regardless of the time of year, the rate of accumulation in summer (July to August) was more than double that of the winter and spring (Table 1): at this time the warmer seawater supported a greater metabolic rate so that more seawater passed over the gills⁸. From July 4 to August 29, the experimental oysters accumulated 45% of the cadmium present by the end of 40 weeks (Table 1). Oysters exposed to 0.0005 p.p.m. cadmium for 40 weeks from November 19 to

August 29 accumulated a mean of 10.75 µg cadmium per g wet weight (Table 1).

Although we have observed that apparently unpolluted seawater contains less than 0.001 p.p.m. of cadmium, values equal to and exceeding 0.01 p.p.m. have been recorded in seawater and freshwater^{9,10} and also in waters from which oysters have been collected¹¹. The high level of cadmium contamination in estuarine waters is exemplified in Corpus Christi Harbor where as much as 0.078 p.p.m. has been reported¹². Therefore, we felt that exposure of oysters to 0.005 p.p.m. cadmium was realistic because such concentrations can be found in the marine environment.

Increased cadmium concentrations seem to be due mainly to trace metal entering estuarine or coastal areas from industrial effluents and rivers. Ratkowsky *et al.*⁶ found a close correlation between proximity to heavily urbanised areas and the concentration of metals in oysters: those growing near urban areas had the greatest concentration of one or more metals. Shuster and Pringle⁵ found concentrations of 0.1–7.80 p.p.m. cadmium (mean 3.10 p.p.m.) in whole wet meats of oysters sampled along the Atlantic coast of the USA. The freshly collected oysters we used contained between 1.78 and 3.78 p.p.m. cadmium wet weight (mean 2.72 p.p.m.).

Table 1 shows that oysters exposed to 0.005 p.p.m. cadmium for 40 weeks accumulated concentrations of cadmium (13–15 p.p.m. wet weight) that could have the same effect on man as foods containing 13–15 p.p.m. (ref 7). The emetic property of cadmium in excess of 13–15 p.p.m. in ingested foods⁷ is a physiological protective mechanism. But concentrations less than 13–15 p.p.m. in foods are also of concern because 5–10% of cadmium ingested by humans is absorbed and retained for many years, chiefly in the kidneys (its biological half life is 16–33 yr)¹. Consequently, cadmium has been implicated as a possible cause of hypertension¹³ through damage to the renal cortex¹.

We thank Dr E. Davey and Mr A. Soper for cadmium analyses.

GERALD E. ZAROOGIAN
SUE CHEER

US Environmental Protection Agency,
Environmental Research Laboratory,
Narragansett, Rhode Island 02882

¹ World Health Organization, *Tech. Rep. Serv.*, 5–32 (1972).

² McKee, J. E., and Wolf, H. W. *Water Quality Criteria*, second ed. Publ. 3-A. (California State Water Quality Control Board, 1963).

³ Shuster, C. N., and Pringle, B. H., *Proc. natn. Shellfisheries Ass.*, 59 (1969).

⁴ Pringle, B. H., Hissong, D. E., Katz, E. L., and Mulawka, S. T., *J. Sanit. Engng Div.*, 94, 355–475 (1968).

⁵ Shuster, C. N., and Pringle, B. H., *Proc. first Mid-Atlantic indust. Waste Conference*, University of Delaware, 285–304 (1968).

⁶ Ratkowsky, D. A., Thrower, S. J., Eustace, I. J., and Olley, J., *J. Fish. Res. Bd Can.*, 31, 1165–1171 (1974).

⁷ Ohio River Valley Water Sanitation Commission, Subcommittee on Toxicities, *Metal Finishing Industries Action Committee Rep. No. 3* (1950).

⁸ Galtsoff, P., *Fish. Bull.*, 64, 397–465 (1964).

⁹ Friberg, L., Piscator, M., and Nordberg, G., *Cadmium in the Environment*, 13–160 (Chemical Rubber Company, Cleveland, 1971).

¹⁰ Elderfield, H., Thornton, L., and Webb, J. S., *Mar. Poll. Bull.*, 2, 44–47 (1971).

- ¹¹ Kopfler, F. C., and Mayer, J., *Proc. Gulf and South Atlantic Shellfish Sanitation Research Conf.*, 1967, 67–80 (US Dept Health, Education, Welfare, 1969).
¹² Holmes, C. W., Slade, E. A., and McLerran, C. J., *Env. Sci. Tech.*, 8 (3), 255–259 (1974).
¹³ Schroeder, H. S., *J. Chronic Dis.*, 18, 647–656 (1956).

Young tomato fruits induced to export carbon by cooling

In phloem translocation systems, sources and sinks are clearly defined: sources are net exporters and sinks net importers of carbon compounds. The overall rate of translocation is largely controlled by ill-defined metabolic processes operating at the source and sink, the transport process itself rarely being limiting^{1–3}. An understanding of how this control is exerted is needed before the mechanism of phloem translocation can be understood, and may provide the potential for manipulating the yield of the economically important parts of plants. The direction of net translocation is reversed as a growing leaf matures^{3,4}, or when previously imported material is redistributed from storage organs such as bulbs^{5,6}. In such cases, however, it is difficult to relate causally the physiology or biochemistry of the organ to the change in the direction of translocation. Previous attempts to convert mature leaves from net exporters to net importers have been only partially successful^{7–10}. We have now achieved complete reversal of the normal role

the experiment the fruit was maintained at either 5 °C (cooled) or 25 °C (control), with the rest of the plant at 25 °C.

Table 1 shows the results of a typical experiment with data from cooled and control fruits grouped according to size. The fruits were rapidly growing and green, and had reached 25–50% of their final size. The smaller control fruits imported carbon at a greater absolute rate than the larger fruits. The amounts of carbon imported in 48 h were 21% and 8% of the initial carbon contents of the smaller and larger fruits, respectively. Cooling greatly reduced the amount of carbon imported by smaller fruits (to 9% of the initial carbon content), but caused a net loss of carbon from larger fruits (by 5% of the initial carbon content); only 7% of this total loss was due to respiration, and so the balance must have been exported from the fruits. Thus the rate and direction of translocation are related to both fruit size and temperature. Sucrose is the major carbon compound translocated in the phloem in the tomato¹¹, but in the green fruit as a whole it comprises less than 2.5% of the dry weight. It is interesting, therefore, that although total sucrose content changed little in the control fruit, there were pronounced increases in the cooled fruits. If the low sucrose content of a green fruit is due to intensive conversion of imported sucrose, the increase in sucrose content of the cooled fruits must be due to either a reduced ability to metabolise imported sucrose or to enhanced synthesis of sucrose from other metabolites.

Table 1 Rates of carbon translocation into (+) or out of (–) tomato fruits at 5 or 25 °C and the concomitant changes in total carbon (Δ carbon) and sucrose (Δ sucrose) during 48 h

Fruit temperature (°C)	Initial carbon content (g)	Δ Carbon (mg per 48 h)	Carbon respired (mg per 48 h)	Carbon translocated (mg per 48 h)	Δ Sucrose (mg per 48 h)
5	1.0–1.3	+ 96.1 ± 15.8	4.5 ± 0.5	+ 100.6 ± 15.4	+ 28.3 ± 7.2
25	1.0–1.4	+ 234.3 ± 10.6	19.1 ± 2.4	+ 253.4 ± 13.0	– 3.4 ± 7.7
5	1.6–1.9	– 84.5 ± 28.8	5.6 ± 0.5	– 78.9 ± 28.8	+ 36.8 ± 8.6
25	2.0–2.2	+ 126.3 ± 16.3	34.5 ± 13.0	+ 160.8 ± 3.8	– 4.8 ± 4.3

Carbon contents were estimated by a dichromate consumption method¹⁴. Respiration was measured by the absorption of CO₂ in NaOH solution. Fruits were cooled by immersion of the gas exchange chambers in ice–water mixtures, thus maintaining the fruit temperatures at 5 ± 1 °C throughout the experiment. Sucrose contents were determined by autoanalysis as the difference between the total sugars and hexoses in 80% ethanol extracts. The results are expressed as the means ± s.e. of two or three fruits.

of a sink, and report here the net export of considerable quantities of carbon from a young, growing tomato fruit.

Carbon translocation to or from an individual fruit was measured as the sum of the change in carbon content and the net respiratory loss of carbon during 48 h. Tomato plants were 'stopped' at the 15th leaf, with all fruits but two on the first truss removed. In this twin-fruit system the carbon contents of the two fruits per unit volume were very similar. Thus the initial carbon content of one fruit could be calculated from its volume if the carbon content and volume of the other fruit were measured. At the start of the experiment the volumes of both fruits were measured; one fruit was then removed for analysis of total carbon, while the other was enclosed in an airtight Perspex chamber for measuring its net loss of CO₂ during 48 h. The final carbon content of this fruit was determined directly. During

The effect of cooling on the import of carbon by the fruit was also examined by estimating amounts of ¹⁴C-leaf assimilate imported during 48 h. The results (Table 2) show that ¹⁴C-leaf assimilate was imported by all the cooled fruits, although to a lesser extent than by the controls. Thus, leaf assimilate may move into a cooled fruit even during periods of net carbon export. Furthermore, the percentage increase of ¹⁴C-sucrose in the larger cooled fruits (found to be net exporters of carbon) was far greater than in the smaller cooled importing fruits, and about six times that in the controls. The greater percentage increase of ¹⁴C-sucrose in the cooled exporters was therefore associated with a high accumulation of sucrose.

Reductions in the accumulation of ¹⁴C-leaf assimilates in the roots and the young leaves of sugar beet¹² and the ear of wheat¹³ have also been achieved by local cooling. In the

Table 2 Percentage increases in total ¹⁴C and ¹⁴C-sucrose in tomato fruits at 5 or 25 °C during 48 h

Fruit temperature (°C)	Initial total carbon content (g)	Net carbon translocation	Total ¹⁴ C (% increase)	¹⁴ C-sucrose (% increase)
5	1.6–1.9	Export	379 ± 34	405 ± 97
5	1.0–1.3	Import	520 ± 21	184 ± 39
25	1.0–1.4	Import	628 ± 29	70 ± 23

¹⁴CO₂ (125 μCi) was applied to the leaf above the truss 24 h before the start of the experiment. The increase in ¹⁴C in the fruit was the difference between its initial and final content of ¹⁴C during 48 h. The initial ¹⁴C content was estimated by the twin-fruit technique described in the text. The results are expressed as means ± s.e. of two or three fruits.

light of our findings there may have been net export of carbon from these cooled sinks as in the tomato, although it could not be detected by the techniques used.

Clearly, actively growing, and therefore importing, fruits can be induced, depending on their size, to export considerable quantities of carbon simply by local cooling. This conversion of a net importer to a net exporter is related to reduced import of leaf assimilate and a high accumulation of sucrose in the cooled fruit. Both the rate and direction of carbon translocation in the tomato seem to be related to changes in the sucrose content of the fruit, not merely to the availability of leaf assimilates. Further experiments involved the injection of ^{14}C -sucrose, ^{14}C -glucose or ^{14}C -fructose individually into fruits which were subsequently cooled for 48 h. In all cases autoradiography of peduncle and stem sections showed that the exported ^{14}C -activity was confined to the phloem.

These findings suggest that the status of an organ as a source or sink is determined by the carbohydrate metabolism of that organ, which may be modified by changes in environmental factors, such as temperature. By modifying the chemical conversions of the unloading process in the sink, the activity of that sink can be altered. The observed temperature effect may to a certain extent explain the observation that tomato fruit growth is reduced at low temperatures even when solar radiation is high.

We thank Dr D. A. Baker for his interest. A.J.W. received an ARC research studentship.

A. J. WALKER
L. C. HO

Plant Physiology Department,
Glasshouse Crops Research Institute,
Littlehampton,
Sussex BN16 3PU, UK

Received February 19; accepted April 27, 1976.

- ¹ Milthorpe, F. L., and Moorby, J., *A. Rev. Pl. Physiol.*, **20**, 117–138 (1969).
- ² Wareing, P. F., and Patrick, J. W., in *Photosynthesis and Productivity in Different Environments*, 3, 481–499 (International Biological Programme, 1975).
- ³ Wardlaw, I. F., *Bot. Rev.*, **34**, 79–105 (1968).
- ⁴ Turgeon, R., and Webb, J. A., *Planta*, **113**, 179–191 (1973).
- ⁵ Ho, L. C., and Rees, A. R., *New Phytol.*, **74**, 421–428 (1975).
- ⁶ Ho, L. C., and Rees, A. R., *New Phytol.*, **76**, 59–68 (1976).
- ⁷ Mason, T. G., and Maskell, E. J., *Ann. Bot.*, **42**, 571–636 (1928).
- ⁸ Mason, T. G., Maskell, E. J., and Phillis, E., *Ann. Bot.*, **50**, 23–58 (1936).
- ⁹ Palmquist, E. M., *Am. J. Bot.*, **25**, 97–105 (1938).
- ¹⁰ Thrower, S. L., *New Phytol.*, **73**, 685–687 (1974).
- ¹¹ Ho, L. C., *Rep. Glasshouse Crops Res. Inst.*, **43** (1972).
- ¹² Geiger, D. R., *Pl. Physiol.*, **41**, 1667–1672 (1966).
- ¹³ Wardlaw, I. F., in *Mechanisms of Regulation of Plant Growth*, Bulletin 12, 533–538 (The Royal Society of New Zealand, Wellington, 1974).
- ¹⁴ Ho, L. C., *J. exp. Bot.*, **27**, 87–97 (1976).

Circadian activity rhythm influenced by near zero magnetic field

CIRCADIAN periodicity has been observed in many organisms. Motor activity, body temperature, and oxygen consumption are among the functions shown to possess a diurnal rhythm. Many of these rhythms are easily entrained to a natural or artificial cycle of light and darkness. In constant laboratory conditions of light and temperature, rhythms persist, but are slightly longer or slightly shorter than 24 h (ref. 1). The persistence of rhythms in constant conditions together with the evidence from translocation experiments², are considered evidence that rhythms are endogenous. Many researchers, among them Pittendrigh³, Aschoff⁴, and Enright⁵, adhere to the endogenous rhythm hypothesis.

The possibility that circadian rhythms might be influenced by some geophysical variable such as the diurnal fluctuations in the intensity of the Earth's normal magnetic field, was suggested by Brown⁶. Brown *et al.*^{6–8} showed that the circadian rhythms of fiddler crabs and other organisms are influenced by small changes in the intensity of the Earth's magnetic field. Wever⁹ showed that the circadian

activity rhythm of humans may be entrained by an artificial 10 Hz electromagnetic field. Keeton *et al.*¹⁰ demonstrated that the orientation ability of pigeons is influenced by small variations in magnetic field intensity. Elimination of the vertical component of the Earth's field influences the orientation ability of the European robin, *Erithacus rubela*¹¹.

Since a time sense is important for orientation, and some evidence suggests that birds can be affected by small changes in artificial magnetic field intensity, we studied the effect of these changes on the circadian rhythm of birds, and have found that the circadian activity of the house sparrow (*Passer domesticus*) can be entrained to a cycle of change in the intensity of the Earth's magnetic field.

The characteristics of the circadian activity rhythm in the house sparrow have been well established by Menaker¹² and Eskin¹³. House sparrows entrain especially well to any reasonable light:dark (LD) cycle. Their period length generally becomes slightly longer in conditions of constant darkness.

The birds were captured in the vicinity of Kingston, Rhode Island. Eight adult males and eight females were selected at random. They were housed singly in non-magnetic 0.028-m³ cages. Two 1.83 m × 1.83 m × 1.37 m Weber coils were constructed with were identical except that only the experimental coil was activated by a Sorenson power supply. The activated coil produced a field with a total intensity of 0 ± 40 nT within a sphere 0.91 m in diameter within the coil. The total intensity of the Earth's field within the control coil was 0.42 gauss, as measured both when the experimental coil was activated, and when it was not activated. A platform was built for each coil so that the cages could be arranged symmetrically within the sphere of 0 gauss. Eight birds (four males and four females) were placed in each coil. Perch hopping activity was recorded following Eskin's method¹³. The wooden perch was long enough to extend outside the boundary of the coil where it was connected to a micro-switch. A barrier was placed parallel to the perch inside the cage, so that the bird would have to hop on the perch to cross the cage. The range of temperature fluctuation was approximately 5 °C. Extraneous noise was masked by a white noise generator which maintained a constant background noise of 84 dB.

For the first 2 weeks all the birds were maintained in an LD 8:16 cycle, and in the Earth's normal magnetic field (MF). For the second 2 weeks all the birds were maintained in the LD 8:16 cycle. The control birds were kept in the Earth's MF, but the experimental birds were exposed to 8 h of near 0 gauss MF and 16 h of the Earth's MF. The 8 h of near 0 gauss MF were synchronised with the 8 h of light. For the last 4 weeks all the birds were maintained in DD. The controls were kept in the Earth's MF. The experimental birds were kept in the MF cycle as described for the previous 2 weeks. The birds were fed and watered, and the intensity of the MF was measured once or twice a week. The field fluctuated only ± 200 nT during the course of the experiment.

The first 2 weeks were used to entrain the birds to the LD 8:16 cycle. The second 2 weeks allowed the experimental birds to adjust to the MF cycle. During the last 4 weeks of the experiment, it was expected that the control birds would free run. It was expected that the experimental birds would show a lesser tendency to free run, if they were in fact capable of being influenced by the MF cycle.

Enright's⁵ periodogram and analysis of covariance¹⁴ were used to test the null hypothesis that there was no difference in the circadian rhythm of the two groups during the 4 weeks in total darkness.

The event recorder records for the first 4 weeks of the

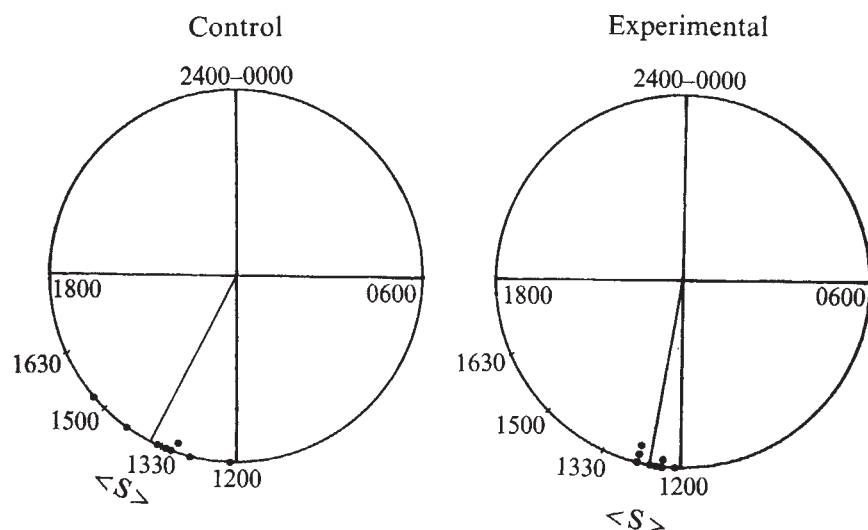


Fig. 1 Mean midpoint of activity in total darkness. This is a circular representation of the times of the midpoints of activity for each bird. 0000 and 2400=360°, 1200=180°. R is a measure of concentration about the mean time (the higher the concentration, the closer R is to 1). S is the mean angular deviation (similar to the standard deviation). The mean time of the midpoint for the control group was 1348 (207°). $R=0.95$. $S=1$ h 7 min (16.75°). The mean time of the midpoint for the experimental group was 1242 (190.5°). $R=0.99$. $S=15$ min (3.75°).

experiment show that both the control birds and the experimental birds entrained strongly to the LD 8:16 cycle as expected. The introduction of an 8:16 MF cycle during the second 2 weeks had no observable effect on the experimental birds.

During the final 4 weeks of the experiment, the period length of most of the control birds became longer; the onset of activity began later and later each day. This resulted in a large amount of activity from 1800 to 2400.

The activity period of the experimental birds lengthened in most cases. This lengthening was, however, due to the activity beginning earlier and ending later than it had while the light cycle was in operation. The event recorder records show no evidence of a phase shift taking place; the midpoint of activity fluctuated about a mean of 12 noon as it had during the light cycle. The activity between 1800 and 2400 was considerably less than that of the controls.

Enright's periodogram was used to determine the total period length. All the birds showed a marked 24-h period length during the first 4 weeks. The control birds had a period length of approximately 24 h 30 min during the last 2 weeks of the experiment. The experimental birds had a period length of 24 h; however, the relative amplitude of the 24-h period was less marked than in LD 8:16. The activity between 1800 and 2400 was significantly greater for the controls than for the experimentals during the 4-week total darkness period (analysis of covariance, $P<0.05$).

Batschelet's¹⁵ method was used to determine the mean times of the midpoint of activity for each bird, both in LD 8:16 and in total darkness. The mean times were determined for the control group, and for the experimental group. In LD 8:16 the F test showed that there was no significant difference between the two groups. In total darkness the time of midpoint of activity for the experimental group was significantly earlier than for the controls ($P<0.025$) (see Fig. 1).

The results demonstrate that the circadian activity of the house sparrow may be entrained by a cycle of change in the intensity of the vertical component of the Earth's MF. The MF seems to act as a weak Zeitgeber compared with the effect of light. House sparrows, however, show a strong degree of synchronisation with a light cycle.

Enright's periodogram analysis shows that the period length of the experimental birds does not become significantly longer in total darkness, from that of the control birds. The activity period of the experimental birds did not shift significantly out of phase with the previous light period, whereas the activity of the control birds did show a shift.

There were no differences in activity of the controls or experimentals during the first 4 weeks. This is probably

because the MF cycle was synchronised with the light cycle, and the 0 MF took the place of light as a potential Zeitgeber only when the light cycle was removed. The experimental birds did not adapt to the MF cycle (in the sense of no longer being influenced by it) during the 6 weeks that the experiment was in operation.

It is possible that the activity of each bird might not be independent of the activity of the other birds. Auditory cues were, however, masked by the white noise generator. Since the cages were placed so that only two (one control bird and one experimental bird) of the birds were able to see each other, visual cues seem an unlikely source of entrainment.

Possibly the ability to be influenced by a near 0 gauss MF is species specific. Or it may influence only such parameters as circadian rhythms and navigation. This might explain why researchers have had little success in demonstrating the effects of intensity changes on human reaction time¹⁶, and other parameters of human behaviour¹⁷; or in training birds to respond to changes in intensity as a conditioned stimulus^{18,19}.

We thank John Haffner, Tom Keefe, Harold Pomeroy, and Eric Scully for their assistance. This work was supported by a private patron.

VIRGINIA L. BLISS
FRANK H. HEPPNER

Department of Zoology,
University of Rhode Island,
Kingston, Rhode Island 02881

Received December 22, 1975; accepted February 12, 1976.

- ¹ Palmer, J. D., *Comp. Biochem. Physiol.*, **12**, 273-282 (1964).
- ² Renner, M., *Cold Spring Harb. Symp. quant. Biol.*, **25**, 361-367 (1960).
- ³ Pittendrigh, C. S., *Cold Spring Harb. Symp. quant. Biol.*, **25**, 159-184 (1960).
- ⁴ Aschoff, J., *A. Rev. Physiol.*, **25**, 581-600 (1963).
- ⁵ Enright, T. J., *J. theor. Biol.*, **8**, 426-468 (1965).
- ⁶ Brown, F. A., Jr., *Cold Spring Harb. Symp. quant. Biol.*, **25**, 57-71 (1960).
- ⁷ Brown, F. A., Jr., in *Circadian Clocks* (edit. by Aschoff, J.), 231-261 (North-Holland, Amsterdam, 1965).
- ⁸ Brown, F. A., Jr., Barnwell, F. H., and Webb, M., *Biol. Bull.*, **127**, 221-231 (1964).
- ⁹ Wever, R., *Naturwissenschaften*, **55**, 29-32 (1968).
- ¹⁰ Keeton, W. T., Larkin, T. S., and Windsor, D. M., *Z. vergl. Physiol.*, **95**, 95-105 (1974).
- ¹¹ Wilschko, W., and Wilschko, R., *Science*, **176**, 62-64 (1972).
- ¹² Menaker, M., in *Circadian Clocks* (edit. by Aschoff, J.), 385-395 (North-Holland, Amsterdam, 1965).
- ¹³ Eskin, A., thesis, Univ. Texas at Austin (1969).
- ¹⁴ Snedecor, G. W., and Cochran, W. G., *Statistical Methods* (Iowa State University, Iowa, 1967).
- ¹⁵ Batschelet, E., *Statistical Methods For the Analysis of Problems in Animal Orientation and Certain Biological Rhythms* (American Institute of Biological Science, Washington DC, 1965).
- ¹⁶ Friedman, H., Becker, R. O., and Bachman, C. H., *Nature*, **213**, 949-956 (1967).
- ¹⁷ Beischer, D. E., *Ann. N. Y. Acad. Sci.*, **188**, 324-330 (1971).
- ¹⁸ Meyer, E. M., and Lambe, D. R., *Psychon. Sci.*, **5**, 349-350 (1966).
- ¹⁹ Emlen, S. T., *Anim. Behav.*, **18**, 215-224 (1970).

Observations on nudging cells in culture

CELL behaviour has been studied widely in tissue culture, and in many cases individual cells have been shown to inhibit the locomotion of cells with which they make contact¹. In the embryo, however, cells move mainly in sheets² or streams³⁻⁵ and in the latter case their locomotion does not appear to be contact inhibited⁶. The question, then, is how the behaviour of individual cells is coordinated in these organised morphogenetic movements. While studying isolated deep cells of the fish, *Fundulus heteroclitus*, we have made some observations which relate to this question.

Deep cells of blastulae form hemispherical bulges of the cell surface called blebs. As cell locomotion begins during gastrulation some blebs extend to form lobopodia or flatten to form lamellipodia⁶. In films of deep cells blebbing within the embryo it is often difficult to determine whether cells are in contact, and if so, to what extent. Fortunately, however, deep cells isolated from blastulae in culture behave in the same way as cells in the embryo⁷.

Blastoderms (stage 11½) were isolated and their deep cells were disaggregated mechanically by flushing through a narrow-bore micropipette. This resulted in a suspension of single cells and small clusters which was placed in a deep watch glass containing a simple culture medium⁸. Most of the cells stuck to the glass substratum within 15 min, although they did not flatten but rather adhered over a small portion of the cell surface as *in vivo* in the blastula. This may have been due to the lesser deformability of cells at this stage⁹. They resumed blebbing during this time. A bulge formed which expanded into a bleb in about 5–8 s. Blebs seemed to be laterally restricted and only involved a small area of the cell surface. Such a bleb was eventually resorbed into the cell and after a variable time another bleb formed almost diametrically opposite the position of the first. The positioning did not seem to be random, in contrast to amphibian gastrula cells¹⁰. In no instance, in observations of more than 50 cells, has a bleb resorbed and a new bleb formed in the same position.

Most blastula cells in culture initiated new blebs every 30 s. To see if cells in contact influenced each others' blebbing, two attached blastula cells were observed. Both were 40 µm in diameter and had a broad contact between them so that each cell was flattened along the contact. These cells were observed every 15 s for 5 min and 70% of the time they were both blebbing or not blebbing. Many observations of non-contacting deep cells would be required to substantiate a correlation of blebbing between cells in contact. Instead of this, we have taken advantage of the fact that we can stimulate a cell to bleb by nudging or stroking it with a micropipette or smoothed glass probe controlled by a micromanipulator.

Eight individual blastula cells were observed in culture for 2 min, during which one of the cells did not bleb while the other seven initiated a new bleb every 30 s. These cells were then nudged. The cell which had not blebbed did not form a bleb when nudged nor during the next minute. Six cells blebbed within 10 s of being nudged, three of these within 5 s. One of these cells already had a bleb but a new bleb was initiated 10 s after nudging. The remaining cell also had a bleb but did not initiate a new bleb after nudging when observed for 1 min. The bleb which formed on nudging was always diametrically opposite the point where the cell was touched, or almost so. In no case did the bleb form where the cell was touched.

Touching the cell thus seems to stimulate blebbing. This suggests that the increased number of blebbing deep cells *in vivo* as development proceeds during the blastula stage⁶ may be due to the increased probability of blebbing cells touching each other. We suggest in addition that this sensitivity to touch *in vitro* also operates when cells move in streams *in vivo*, as in the *Fundulus* germ ring during gastrulation, and

that cell movement forward is enhanced by cells bumping into the rear of cells ahead of them. This mechanism would only operate if the cell movement is directional, as in morphogenetic movements. Curtis¹¹ has pointed out that such shearing of cells against each other would effect viscosity changes in their surfaces, and in appropriate conditions could promote motility.

Amoebae are also stimulated to produce pseudopodia when prodded with a needle, but in this case the protrusion is formed near the site of stimulation. Goldacre¹² has suggested that production of a pseudopod is stimulated by the membrane interacting with the endoplasm of the cell. Although in deep cells there is no evidence for sol-gel transformation, the blebs formed may be analogous to the ectoplasm of an amoeba in that cytoplasm seems to flow into a bleb as it elongates to form a longer protrusion.

To test whether the stimulus for blebbing could be transmitted to another cell, nine doublets of blastula cells adherent to the substratum were studied. One of the two cells was chosen at random to be nudged away from the region of contact. In one doublet, neither the cell nudged nor the other blebbed for 30 s after nudging. In all other doublets both cells blebbed simultaneously; five after 6 s, two after 10 s, and one after 25 s. The blebs formed in the nudged cells in the same position as in single cells. The blebs forming in the attached cells were usually at one side of the region of contact between the cells. This suggests that the surface activity of the cells of the doublet is linked.

As *Fundulus* deep cells have been shown to be electrically coupled¹³, we suggest that this might be the mechanism whereby surface activities of cells are coordinated. It is clear, however, that contact with another cell also restricts the activity of an individual cell. Blebs do not form at points of contact and it is possible that the cell surface is not free to flow in these regions (our work in preparation). The question is whether surface activity in groups of cells is coordinated by interaction of the membranes themselves or by messages passing through communicating channels between cells. The same question applies to the local effects on ruffling¹⁴ and contraction¹⁵ on contact with another cell.

This work was supported by grants from the NSF and NIH.

C. TICKLE

Department of Biology as Applied to Medicine,
Middlesex Hospital Medical School,
London W1P 6DB, UK, and
Marine Biological Laboratory,
Woods Hole,

J. P. TRINKAUS

Department of Biology,
Yale University,
New Haven, Connecticut 06520, and
Marine Biological Laboratory,
Woods Hole, Massachusetts 02543

Received March 22; accepted April 20, 1976.

- ¹ Abercrombie, M., *In Vitro*, **6**, 128–142 (1970).
- ² Vogt, W., *Arch. Entw. Organ.*, **120**, 384–706 (1929).
- ³ De Haan, R. L., *Exptl Cell Res.*, **29**, 544–560 (1963).
- ⁴ Weston, J. A., *Adv. Morphogen.*, **8**, 41–108 (1970).
- ⁵ Trinkaus, J. P., *Cells into Organs* (Prentice-Hall, Englewood Cliffs, New Jersey, 1969).
- ⁶ Trinkaus, J. P., *Dev. Biol.*, **30**, 68–103 (1973).
- ⁷ Trinkaus, J. P., *Dev. Biol.*, **7**, 513–532 (1963).
- ⁸ Armstrong, P. B., and Child, J. S., *Biol. Bull.*, **128**, 143–154 (1965).
- ⁹ Tickle, C. A., and Trinkaus, J. P., *J. Cell Sci.*, **13**, 721–726 (1973).
- ¹⁰ Lucy, E. C. A., and Curtis, A. S. G., *Med. Biol. Illust.*, **9**, 86–93 (1959).
- ¹¹ Curtis, A. S. G., *Exptl Cell Res. Suppl.*, **8**, 107–122 (1961).
- ¹² Goldacre, R. V., *Symp. Soc. exp. Biol.*, **6**, 128–144 (1952).
- ¹³ Bennett, M. V. L., and Trinkaus, J. P., *J. Cell. Biol.*, **44**, 592–610 (1970).
- ¹⁴ Trinkaus, J. P., Betchaku, T., and Krulikowski, L. S., *Exptl Cell Res.*, **64**, 291–300 (1971).
- ¹⁵ Abercrombie, M., and Dunn, G., *Exptl Cell Res.*, **92**, 57–62 (1975).

Diminished content of plasma membrane-associated myosin in transformed fibroblasts

MYOSIN, a major contractile element of muscle, has been found in a variety of non-muscle cells such as platelets^{1,2}, fibroblasts^{3,4}, brain tissue⁵, amoebae⁶ and slime mould⁷. It is localised in the plasma membrane⁸ as well as in the cytosol⁴ of cultured fibroblasts. Although no definite function has been established for non-muscle myosins, they may participate in cellular movement, mitosis, changes in cell shape and related functions. We have examined the protein subunit composition of plasma membranes from normal and transformed fibroblasts with a particular emphasis on determining the content of plasma membrane-associated myosin. We present here evidence that the content of plasma membrane-associated myosin is diminished in transformed fibroblasts. This decrease may in part account for the abnormal shape and loss of contact inhibition of movement of transformed cells⁹.

Plasma membranes from normal and transformed rat kidney fibroblasts were prepared by a slight modification of the method of Hochstadt *et al.*¹⁰ in which subcellular fraction is carried out using hypertonic sucrose and its polymer Ficoll. Examination of the membrane fraction by electron microscopy showed that about 90% of the structures were closed membrane vesicles and less than 10% were mitochondria. Marker enzyme studies using adenylate cyclase and Na⁺, K⁺-ATPase confirmed that the membrane fraction contained purified plasma membrane, with enrichment ranging up to 10 times greater than homogenates. Furthermore, the contamination of the membrane fraction with endoplasmic reticulum, as assessed by the specific activity of the marker enzyme NADH dehydrogenase¹⁰, was comparable between normal and transformed cell lines.

Figure 1 shows a typical electrophoretogram of homogenates and plasma membranes from normal and transformed cells, in which equal amounts of protein were

applied. In this experiment, the protein subunit patterns from the normal and transformed whole cell homogenates were quite similar. On the other hand, when the protein subunit patterns of plasma membranes of transformed fibroblasts were compared with those of normal fibroblasts, a marked decrease was observed in a protein band with a molecular weight of 200,000. This band corresponds to the heavy chain of fibroblast myosin as indicated by the arrows in Fig. 1. As previously noted the heavy chain of fibroblast myosin migrates slightly faster than the heavy chain of the skeletal muscle myosin standard used^{3,4}. The band in the plasma membrane preparations migrating slightly slower than myosin probably represents the myosin-associated

Table 1 Content of membrane-associated 200,000-molecular weight protein

Cell lines	µg myosin-mg membrane protein	%
NRK	42	100
HNRK	19	45
KNRK	13	31
MNRK	11	26
SNRK	16	38

Quantitative determination was performed using Gilford model 400 spectrophotometer equipped with a densitometric apparatus and slab electrophoretic gels with three different preparations of normal and transformed membranes. The content of the protein in the membranes was computed by dividing the area of protein in the myosin band (determined by densitometry) by the total area of all the stained bands, and the results are expressed as µg per mg of total membrane protein.

protein of molecular weight 210,000, recently described by Olden *et al.*¹¹, and is immunologically unrelated to myosin (K.O., M. C. Willingham and I.P., manuscript in preparation). The quantitative measurement of the myosin band by densitometry using purified myosins from L cells and rabbit muscles as standard calibration proteins indicated that myosin was decreased by 60–75% in transformed membrane preparations. Table 1 summarises the average values of membrane-associated protein of molecular weight 200,000 from normal and transformed cells when three different plasma membrane preparations were used.

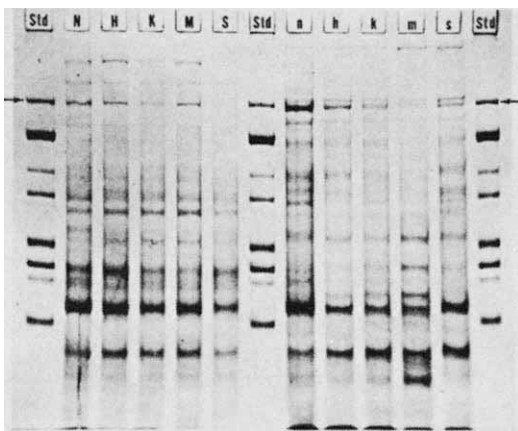


Fig. 1 Normal rat kidney fibroblasts (NRK), and NRK cells transformed with Harvey (HNRK), Kirsten (KNRK), Moloney (MNRK) and Schmidt-Ruppin (SNRK) viruses were grown to confluency in 32-ounce bottles using Dulbecco-Vogt medium supplemented with 10% calf serum. Each bottle was washed twice with 100 ml of 10 mM phosphate buffer, pH 7.5, containing 0.15 M NaCl (PBS) and once with 100 ml of 10 mM Tris-HCl, pH 8.0, containing 0.5 M sucrose, 0.1 mM dithiothreitol and 0.1 mM EDTA (Tris-sucrose buffer). The cells were scraped into 10 ml of buffer. The plasma membrane fraction was isolated at 1–4 °C by the following modification of the method of Hochstadt *et al.*¹⁰. The mixture from each bottle was homogenised with 10 strokes of a Potter homogeniser and unbroken cells were further homogenised with five strokes of a Dounce homogeniser. A

sample was taken from each homogenate for electrophoresis. Then 5 µg of DNase and magnesium sulphate at final concentration of 0.1 mM were added to the residual homogenates. The nuclear fraction was removed by centrifugation at 1,000g for 5 min. The supernatants were further centrifuged in a SW30 rotor at 4 °C at 20,000g for 15 min. Each pellet thus obtained was suspended in 5 ml of Tris-sucrose buffer and the suspension layered on 10% Ficoll dissolved in Tris-sucrose buffer. Centrifugation using a Beckman SW40 rotor was performed at 100,000g for 2 h. The interface layer was collected and washed with 10 ml of Tris-sucrose buffer containing 0.1 mM magnesium sulphate, followed by centrifugation at 20,000g for 15 min. The pellet was suspended in a small amount of Tris-sucrose buffer containing 0.1 mM magnesium sulphate and designated as the plasma membrane fractions. Then, each membrane fraction or whole cell homogenate from normal and transformed cells was dissolved in 2% SDS containing 100 mM dithiothreitol, 10 mM potassium phosphate buffer, pH 7.0, 29% glycerol and 0.03% bromophenol blue. Samples were heated at 100 °C for 3 min. Approximately 25 µg of each sample was applied on a SDS slab 5% polyacrylamide gel (1 mm width) and electrophoresis was performed according to the method of Studier¹². Protein staining was performed using 0.25% Coomassie blue in 50% trichloroacetic acid. Excess dye was removed by incubating the gel with 7% acetic acid. Std represents the gel electrophoretic pattern of the following molecular weight standard proteins: skeletal muscle myosin (200,000, see arrow), RNA polymerase (160,000 and 150,000), β-galactosidase (130,000), phosphorylase a (95,000), bovine serum albumin (68,000) and ovalbumin (43,000). On the left are the gel patterns of whole cell homogenate from NRK (N), HNRK (H), KNRK (K), MNRK (M) and SNRK (S), and on the right are the gel patterns of the membrane fractions from NRK (n), HNRK (h), KNRK (k), MNRK (m) and SNRK (s). The lower band of the doublet in the 200,000-molecular weight range of the membrane fraction is myosin. The upper band correlates to a myosin-associated protein¹¹.

Table 2 EDTA-ATPase activity in whole cell homogenates and membrane preparations from normal and transformed cells

Cell lines	EDTA-ATPase			
	Homogenate (nmol min ⁻¹ mg ⁻¹)	%	Membrane (nmol min ⁻¹ mg ⁻¹)	%
NRK	0.88	100	1.41	100
HNRK	0.63	72	0.57	40
KNRK	0.67	76	0.47	33
MNRK	0.68	77	0.52	37
SNRK	0.73	83	0.58	41

EDTA-ATPase activity was determined by the following modification of the method of Ostlund *et al.*⁴. The reaction mixture containing 25 mM 2-N-morpholinoethane sulphonic acid buffer, pH 7.0, 0.2 mM γ -³²P-ATP (50 c.p.m. pmol⁻¹), 0.6 M KCl, 2 mM EDTA, 1 mM dithiothreitol, 1 mM ouabain and the enzyme sample in a final volume of 0.1 ml. Incubation was performed at 37 °C for 20 min. The reaction was stopped by adding 0.5 ml of 20% trichloroacetic acid and the precipitate formed was removed by centrifugation. To 3 ml of supernatant were added 0.1 ml each of 5% ammonium molybdate in 5 N sulphuric acid and 30 mM silicotungstic acid. The molybdate-phosphate complex was extracted with 2 ml of 50% isobutyl alcohol and 50% benzene¹³, and sample of the organic phase was used for determining radioactivity. This assay was linear with respect to time and amount of the enzyme sample in the conditions used.

To identify the 200,000-molecular weight protein band as the heavy chain of myosin, the sodium dodecyl sulphate (SDS) gels were stained as before (K.O. and K. M. Yamada, manuscript in preparation), using goat antimyosin antibody, rabbit antigoat λ -globulin conjugated with horseradish peroxidase, 3,3'-diaminobenzidine and hydrogen peroxide. A peroxidase-positive band was observed in the position of the 200,000-molecular weight protein, indicating that the protein that was decreased in the transformed cell membranes was immunochemically very similar to fibroblast myosin.

To confirm that the content of membrane-associated myosin was diminished in transformed fibroblasts, we measured the characteristic EDTA-ATPase activity (myosin ATPase)³ using normal and transformed cell membranes as well as whole cell homogenates. Table 2 shows the results of such experiments. The levels of EDTA-ATPase activity in the transformed cell plasma membranes were always 20–40% of that in the normal cell membranes, whereas the EDTA-ATPase activities of the whole cell homogenates from the transformed cells were always 70–95% of that in the normal cell homogenate. Another point to be noted in Table 2 is that the concentration of myosin in the plasma membrane fraction is only slightly greater than its concentration in the whole cell homogenate. Thus myosin is not concentrated in the plasma membrane fraction.

Table 3 shows the effect of antimyosin λ -globulin on EDTA-ATPase activity of purified fibroblast myosin and a fibroblast membrane preparation. In the conditions used approximately 60% of EDTA-ATPase activity was inhibited with antimyosin λ -globulin when both purified myosin and fibroblast membranes were used. Control globulin preparations inhibited EDTA-ATPase activity by less than 15%, indicating that the bulk of EDTA-ATPase

activity measured is derived from myosin ATPase activity. Since the results obtained from the quantitative densitometric analyses of membrane-associated myosin (Table 1) agree with the data showing the decreased levels of EDTA-ATPase activity (Table 2), we conclude that membrane-associated myosin is diminished in these transformed rat kidney fibroblasts.

Recent reports from several laboratories^{14–16} have indicated that the content of 200,000-molecular weight protein subunit is decreased in plasma membranes from transformed chick embryo fibroblasts. Although the 200,000-molecular weight protein was not identified, it seems likely that the protein subunit that is diminished in chick cell membranes is also the heavy chain of myosin, if the present and previous experimental results^{4,8} from our laboratory are compared with these reports^{14–16}. It needs to be emphasised that myosin is distinct from a higher molecular weight glycoprotein (220,000) recently isolated from the surface of chick embryo cells¹⁷ with a content markedly decreased after transformation or trypsin treatment.

Treatment of some transformed cells with dibutyryl cyclic AMP induces flattening and elongation, morphological changes which make the transformed cells resemble in some respects their normal parent cells (see review in ref. 18). Therefore we attempted to determine the content of membrane-associated myosin as well as EDTA-ATPase activity in plasma membranes from a SV40-transformed BALB/3T3 cell line that was cultured in the presence and absence of 1.5 mM dibutyryl cyclic AMP and 0.5 mM methyl isobutyl xanthine, a phosphodiesterase inhibitor. This cell rather than a transformed NRK cell was chosen because it has a very striking morphological response to the nucleotide. The results indicated the EDTA-ATPase activity and the content of membrane-associated myosin was increased by less than 30% when the cells were treated with dibutyryl cyclic AMP. Thus the amount of myosin associated with the plasma membrane in this cell line does not seem to be affected by cyclic AMP. The mechanism responsible for the low content of myosin in plasma membranes of these transformed cells is under investigation.

We thank Dr M. C. Willingham for taking electron microscopic photographs and Ms E. Lovelace for technical assistance.

YUTAKA SHIZUTA
PETER J. A. DAVIES
KENNETH OLDEN
IRA PASTAN

Laboratory of Molecular Biology,
National Cancer Institute,
National Institutes of Health,
Bethesda, Maryland 20014

Received January 13; accepted April 20, 1976.

- Adelstein, R. S., Pollard, T. D., and Kuehl, W. M., *Proc. natn. Acad. Sci. U.S.A.*, **68**, 2703–2707 (1971).
- Bettex-Galland, M., and Luescher, E. F., *Adv. Protein Chem.*, **20**, 1–35 (1965).
- Adelstein, R. S., Conti, M. A., Johnson, G. S., Pastan, I., and Pollard, T. D., *Proc. natn. Acad. Sci. U.S.A.*, **69**, 3693–3697 (1972).
- Ostlund, R. E., Pastan, I., and Adelstein, R. S., *J. biol. Chem.*, **249**, 3903–3907 (1974).
- Berl, S., Puszkun, S., and Nicklas, W. J., *Science*, **179**, 441–446 (1973).
- Pollard, T. D., and Korn, E. D., *J. biol. Chem.*, **248**, 4682–4690 (1973).
- Hatano, S., and Ohnuma, J., *Biochim. biophys. Acta*, **205**, 110–120 (1970).
- Willingham, M. C., Ostlund, R. E., and Pastan, I., *Proc. natn. Acad. Sci. U.S.A.*, **71**, 4144–4148 (1974).
- Abercrombie, M., *In Vitro*, **6**, 128–142 (1970).
- Hochstadt, J., Quinlan, D. C., Rader, R. R., Li, C.-C., and Dowd, D., in *Methods in Membrane Biology*, **5** (edit. by Korn, E.), 117–162 (Plenum, New York, 1975).
- Olden, K., Willingham, M. C., and Pastan, I., *J. Cell Biol.*, **67**, 316a (1975).
- Studier, F. W., *J. molec. Biol.*, **79**, 237–248 (1973).
- Martin, J. B., and Doty, D. M., *Analyt. Chem.*, **21**, 965–967 (1949).
- Stone, K. R., Smith, R. E., and Joklik, W. K., *Virology*, **58**, 86–100 (1974).
- Marciani, D. J., and Bader, J. P., *Biochim. biophys. Acta*, **401**, 386–398 (1975).
- Isaka, T., Yoshida, M., Owada, M., and Toyoshima, K., *Virology*, **65**, 226–237 (1975).
- Yamada, K. M., Yamada, S., and Pastan, I., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 3158–3162 (1975).
- Pastan, I., Johnson, G. S., and Anderson, W. B., *A. Rev. Biochem.*, **44**, 491–522 (1975).

Table 3 Effect of antimyosin γ -globulin on EDTA-ATPase activity

Enzyme source	EDTA-ATPase (nmol min ⁻¹ mg ⁻¹)	Inhibition (%)
Myosin from L cells	6.0	0
+ antimyosin γ -globulin (100 μ g)	2.1	65
+ antimyosin γ -globulin (200 μ g)	2.2	63
Membrane preparation from NRK	1.48	0
+ antimyosin γ -globulin (100 μ g)	0.54	64
+ antimyosin γ -globulin (200 μ g)	0.58	61
+ antirabbit γ -globulin (200 μ g)	1.26	15
+ human γ -globulin (200 μ g)	1.36	8

The assays were performed as described in the legend to Table 2. Purified L-cell (15 μ g) myosin and 100 μ g of NRK membrane protein were used.

Two mechanisms for spontaneous recovery from depolarising drugs in rat muscle

THE action of depolarising drugs on the endplate of skeletal muscle is difficult to interpret because the depolarisation is often transient and followed by spontaneous recovery during the continued presence of the drug, and because of differences between species. In frog muscle, depolarising drugs produce an initial increase in conductance at the endplate, followed by a decline in the conductance as 'desensitisation' develops¹⁻³. In cat muscle, depolarisation is maintained while the drug is applied⁴. Rat muscle has long been considered anomalous in that the depolarisation is transient and limited⁵. We now report an unexpected finding for rat diaphragm, in which the recovery is attributed to the action of an electrogenic sodium pump⁶, stimulated by the entry of sodium.

Figure 1 shows the action of carbachol (100 μ M) recorded with an internal electrode with tetrodotoxin (0.1 μ M) in the outside bath to prevent action potentials. The recovery of the membrane potential in the presence of the drug might be due to closure of the ion channels after the initial action of the drug, or to another process. There are methods by which the entry of labelled sodium in the junctional region can be studied directly in diaphragm muscle⁷, and dose-response curves have been obtained⁸. Figure 2a shows the uptake of ²⁴Na in a diaphragm exposed for 15 s to a solution containing ²⁴Na and carbachol

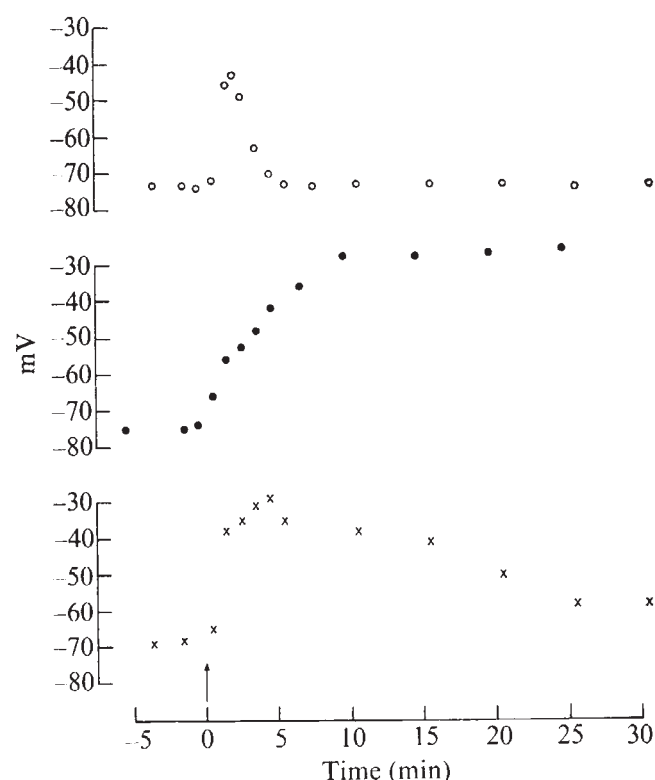


Fig. 1 ○, Effect of 100 μ M carbachol at endplate, with internal recording¹⁹; muscle vertical, arranged for continuous register and rapid changeover, at 38 °C. The solution contained 5 mM potassium and 0.1 μ M tetrodotoxin. The ordinate gives membrane potential, with depolarisation upwards. The abscissa gives time after the solution was changed. Miniature endplate potentials disappeared when carbachol was applied. ●, Effects of 100 μ M carbachol plus 100 μ M ouabain. The reading after 25–60 min with ouabain plus carbachol 100 μ M was –31 mV (median of six estimations). x, Effect of 3 mM decamethonium plus 100 μ M ouabain: final reading, compared with initial resting potential, showed depolarisation of 9 mV (median of nine estimations).

(100 μ M) followed by washing for 6 min. Compared with the control (Fig. 2b), there was an extra uptake of ²⁴Na in the junctional region in the presence of the depolarising drug. The effect was similar (Fig. 2c) when the diaphragm was pretreated with carbachol (100 μ M) for 30 min. These results (and others to be reported elsewhere) are consistent with the view that the ion channels continue to open while the drug is in contact with the tissue, and this differs markedly from the situation in frog muscle.

The recovery of the resting potential could be attributable to an electrogenic sodium pump, and in this case the process would be affected by certain enzyme inhibitors. When ouabain⁹ and carbachol (100 μ M) were added concomitantly (Fig. 1) there was a profound depolarisation and little evidence of recovery. In rat diaphragm, ouabain alone (100 μ M at 38 °C) produced a depolarisation of 9 mV (median of 10 estimations, range 6–17 mV) by sampling methods used before¹⁰. This is a greater effect than that found in frog muscle at room temperature with strophanthidin¹¹.

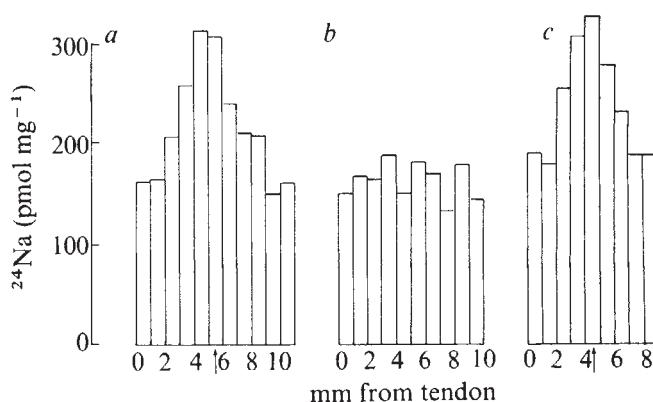


Fig. 2 a, Effect of carbachol on influx of labelled sodium in rat diaphragm in the presence of 0.1 μ M tetrodotoxin. The diaphragm was exposed for 15 s to solution containing carbachol (100 μ M) plus ²⁴Na, (1 mCi in 10 ml), then washed for 6 min in inactive saline, frozen, sliced and counted^{7,20}. Ordinate gives ²⁴Na (pmol mg⁻¹); abscissa is distance from tendon (mm). The arrow indicates the strip of muscle which contained the band of endplates²⁰. b, Control without carbachol. The uptake in a, above the level assigned to the ends of the muscle, gives the extra sodium which entered in the presence of carbachol and which remained after washing for 6 min: extrapolation to zero time is needed to give the uptake at the end of the exposure. c, Uptake in a diaphragm which had been incubated for 30 min with 100 μ M carbachol and then treated as in a.

Figure 1 shows that the membrane potential returned, in the presence of carbachol, to a level close to the initial reading, and in nine experiments the difference ranged from –9 mV to +4 mV. It is not clear how the voltage could be regulated to give recovery to the previous value, and in other situations, with a higher initial resting potential, the recovery from depolarising drugs was incomplete in rat muscle⁵.

Decamethonium (100 μ M) in the presence of ouabain had an effect similar to that shown in Fig. 1 (closed circles). But with large concentrations of decamethonium the situation was different. Figure 3a shows the uptake of ²⁴Na in a muscle exposed for 15 s to decamethonium (3 mM) and washed for 6 min. There was an extra uptake in the junctional region compared with the value at the ends of the fibres, and this initial entry was comparable with the effect shown in Fig. 2a. Pretreatment with decamethonium (3 mM) for 30 min (Fig. 3b) gave little or no indication of an extra uptake of sodium in the junctional region, and we infer that after 30 min at this concentration the channels for sodium had mostly been closed. When a solution containing ouabain (100 μ M) and decamethonium (3 mM) was applied, there was an initial depolarisation (Fig. 1, crosses) followed by marked spontaneous recovery.

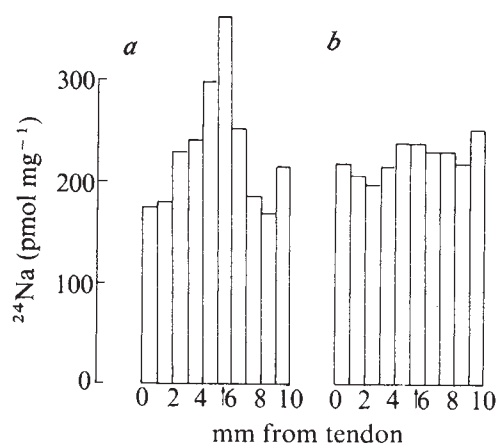


Fig. 3 Effect of 3 mM decamethonium on entry of ^{24}Na : exposure for 15 s, followed by washing for 6 min and treatment as in Fig. 2. The arrow shows the slice of muscle which contained the band of endplates. In *b* the muscle was incubated for 30 min with 3 mM decamethonium and then treated as in *a*.

In rat muscle there seems to be at least two mechanisms for the recovery of the resting potential. The first is an ouabain-sensitive process¹² in which sodium channels continue to open under the influence of the drug. This may explain why labelled decamethonium, which enters at the endplate region¹³, accumulates steadily for many hours¹⁴ in spite of desensitisation. The second process is seen with high concentrations of decamethonium and is largely unaffected by ouabain, apparently involving closure of sodium channels. The permeability to decamethonium falls in the mM range in certain conditions¹⁵, and this could be due to channel block as postulated for frog muscle³. The alternative explanation in terms of receptor mechanisms is that the pharmacological receptors can be transformed to a desensitized state¹⁶⁻¹⁸ in which ion channels become secondarily closed.

This research was supported by the MRC.

R. CREESE
G. I. FRANKLIN
L. D. MITCHELL

Department of Physiology,
St Mary's Hospital Medical School,
London W2 1PG, UK

Received February 26; accepted April 13, 1976.

- 1 Manthey, A. A., *J. gen. Physiol.*, **49**, 963-976 (1966).
- 2 Jenkinson, D. H., and Terrar, D. A., *Br. J. Pharmacol.*, **47**, 363-376 (1973).
- 3 Adams, P. R., *Pflügers Arch. ges. Physiol.*, **360**, 135-144 (1975).
- 4 Head, S. D., *J. Physiol., Lond.*, **241**, 102-103P (1974).
- 5 Thesleff, S., *Acta physiol. scand.*, **34**, 386-392 (1955).
- 6 Thomas, R. C., *Physiol. Rev.*, **52**, 563-594 (1972).
- 7 Creese, R., and Mitchell, L., *J. Physiol., Lond.*, **246**, 44-45P (1975).
- 8 Creese, R., Franklin, G. I., Humphrey, P. P. A., and Mitchell, L. D., *J. Physiol., Lond.*, **254**, 43-44P (1976).
- 9 Scou, J. C., *Biochim. biophys. Acta*, **42**, 6-23 (1960).
- 10 Creese, R., Scholes, N. W., and Taylor, D. B., *J. Pharmac. exp. Ther.*, **124**, 47-52 (1958).
- 11 Terrar, D. A., *Br. J. Pharmacol.*, **51**, 259-268 (1974).
- 12 Mooji, J. J. A., Evers, C. D. J., and Ras, D., *Eur. J. Pharmacol.*, (in the press).
- 13 Creese, R., and MacLagan, J., *J. Physiol., Lond.*, **210**, 363-386 (1970).
- 14 Nedegaard, O. A., and Taylor, D. B., *Experientia*, **22**, 521-522 (1966).
- 15 Creese, R., and England, J. M., *J. Physiol., Lond.*, **210**, 345-361 (1970).
- 16 Katz, B., and Thesleff, S., *J. Physiol., Lond.*, **138**, 63-80 (1957).
- 17 Rang, H. P., and Ritter, J. M., *Molec. Pharmacol.*, **6**, 357-382 (1970).
- 18 Sugiyama, H., and Changeux, J. P., *Eur. J. Biochem.*, **55**, 505-515 (1975).
- 19 Creese, R., Scholes, N. W., and Whalen, W. J., *J. Physiol., Lond.*, **140**, 310-317 (1958).
- 20 England, J. M., *J. Anat.*, **106**, 311-321 (1970).

Psychoactive drug effects on a system which generates cyclic AMP in brain

RECENTLY a direct correlation has been reported between spontaneous motor activity, tyrosine hydroxylase activity and the responsiveness of the noradrenaline sensitive cyclic AMP-generating system in rat brain^{1,2}. These observations

may indicate that the turnover and/or level of noradrenaline in the brain is important in determining the sensitivity of the cyclic AMP-generating system to noradrenaline. Indeed, it has been shown that depletion of brain noradrenaline by the intraventricular injection of 6-hydroxydopamine results in supersensitivity of the system to noradrenaline^{3,4}. Furthermore treatment with reserpine, a drug known to cause depressive reactions in man, has been reported to increase the sensitivity of the adenylate cyclase of rat brain to noradrenaline⁵ and several antidepressant drugs have been shown to have the opposite effect in limbic forebrain⁶. By contrast, we now present evidence that the chronic treatment of rats with both an antidepressant, imipramine, and a depressant, chlorpromazine, reduces the sensitivity of the cyclic AMP-generating system to noradrenaline.

Male Sprague-Dawley rats (initial weight 170-180 g) were used in this study. Treatment was carried out once daily. Rats were given 20 mg kg⁻¹ of either chlorpromazine hydrochloride or imipramine hydrochloride by an oesophageal tube. Control rats received a corresponding volume of water by the same route. Experiments were always performed 24 h after the last drug application. After specified intervals of treatment rats were decapitated and cerebral cortical slices of 260 μm were prepared as rapidly as possible. Procedures for the preparation, incubation and fixation of tissue have been described in detail^{7,8}. Cyclic AMP was determined by the method of Gilman⁹ after purification on Dowex 1 $\times$ 8 columns, each sample being analysed at three different concentrations. All values were corrected for recovery and cyclic AMP values were calculated as pmol mg⁻¹ protein. Statistical comparisons of results were made using Student's *t* test.

No change of the noradrenaline sensitive cyclic AMP formation was observed in control rats treated for 1 to 21 d with water, therefore all values obtained from these animals were combined (Table 1). When rats were treated with either chlorpromazine or imipramine no significant alteration of the noradrenaline elicited accumulation of cyclic AMP was found after 1 and 3 d of drug administration. After 6 d of daily treatment with drugs, however, a reduced sensitivity developed to noradrenaline and cyclic AMP levels were about 40% lower than controls. This reduced sensitivity of the cyclic AMP-generating system did not change during prolonged drug administration for more than two weeks, when cyclic AMP levels only reached 60% of respective controls after incubation with noradrenaline. There was no difference between animals treated with imipramine or chlorpromazine. The changes in the cyclic AMP system appeared after the same period of pretreatment with either drug and were quantitatively undiscernible in spite of the pronounced differences in their therapeutic effects.

The antidepressant imipramine has been shown to inhibit effectively the uptake of noradrenaline and other biogenic amines into peripheral and central monoaminergic neurones. Imipramine may also interfere with the control of noradrenaline release by blocking presynaptic α -adrenoceptors¹⁰. Such potentiation of catecholamines at adrenergic receptor sites is generally assumed to be involved in the mechanism of action of imipramine. If this is so then the data presented here indicate that enhanced stimulation of postsynaptic receptors results in the development of subsensitivity to the neurohormone. The therapeutic action of antidepressant drugs occurs only after chronic treatment, so it is tempting to speculate that the decrease of receptor function is connected with the drugs' therapeutic action⁶.

The mechanism of action of chlorpromazine has been discussed mostly in terms of a blockade of central dopaminergic receptors which is apparently common to nearly all antipsychotic drugs^{11,12}. Therefore, it was surprising to find identical results with chlorpromazine pretreatment on the noradrenaline sensitive adenylate cyclase of

Table 1 Effect of imipramine and chlorpromazine treatment on noradrenaline stimulated formation of cyclic AMP in rat brain cortex

Kind of treatment	Duration of treatment (d)	Cyclic AMP (pmol per mg protein $\pm$ s.e.m.)		
		None	Additions	Noradrenaline, 100 μ M
Water	1-21	13.6 $\pm$ 0.8 (35)		54.6 $\pm$ 3 (35)
Chlorpromazine (20 mg kg ⁻¹)	1	14.0 $\pm$ 1.8 (5)		43.8 $\pm$ 2.9 (5)
Imipramine (20 mg kg ⁻¹)	1	13.8 $\pm$ 3.1 (5)		48.6 $\pm$ 5.2 (5)
Chlorpromazine	3	18.8 $\pm$ 6.7 (5)		50.0 $\pm$ 7.9 (5)
Imipramine	3	14.6 $\pm$ 1.4 (5)		41.0 $\pm$ 7.2 (5)
Chlorpromazine	6	10.6 $\pm$ 2.2 (5)		32.4 $\pm$ 4.5 (5)*
Imipramine	6	11.0 $\pm$ 1.5 (5)		27.0 $\pm$ 4.0 (5)†
Chlorpromazine	10	10.8 $\pm$ 2.5 (5)		33.8 $\pm$ 4.1 (5)*
Imipramine	10	10.0 $\pm$ 0.5 (5)*		32.6 $\pm$ 5.5 (5)*
Chlorpromazine	15	18.4 $\pm$ 2.0 (5)		33.2 $\pm$ 6.5 (5)*
Imipramine	15	15.3 $\pm$ 1.8 (8)		34.5 $\pm$ 4.4 (8)*
Chlorpromazine	21	16.3 $\pm$ 2.8 (8)		35.3 $\pm$ 2.3 (7)*
Imipramine	21	15.2 $\pm$ 3.1 (5)		31.8 $\pm$ 5.8 (5)†

Last drug administration was always 24 h before the experiments. Numbers in brackets refer to individual experiments, not multiple determinations. * $P < 0.025$, † $P < 0.015$, ‡ $P < 0.005$.

rat cerebral cortical slices. Chlorpromazine has, however, been shown to increase the stimulation-induced noradrenaline overflow from field stimulated rat brain slices^{13,14}, resulting in an enhanced availability of catecholamine at postsynaptic receptor sites, although in some systems it is less active than imipramine in blocking noradrenaline reuptake. The fact that chlorpromazine also blocks α -adrenoceptors, possibly presynaptically located, may also contribute to the enhanced presence of catecholamines at the synaptic cleft^{13,15}. This may then trigger the development of subsensitivity of adrenergic receptors to noradrenaline as measured in this study by the formation of cyclic AMP. Any subsensitivity due to a postsynaptic block of α -receptors might be expected to occur after the first dose of chlorpromazine or imipramine. Because of the different therapeutic actions of imipramine and chlorpromazine the significance of the reduced sensitivity of the cyclic AMP-generating system to noradrenaline as a possible common feature of antidepressants requires re-evaluation. It must, however, be remembered that a drug may have more than one action and the fact that chlorpromazine has been shown to have an effect previously only associated with antidepressants does not completely invalidate such an association. Indeed the value of chlorpromazine in conditions like schizophrenia could well depend on its potential antidepressant properties to counter marked depression.

The technical assistance of Ms G. Kleefeld is gratefully acknowledged. This study was supported by the Deutsche Forschungsgemeinschaft.

JOACHIM SCHULTZ

Pharmazeutisches Institut der Universität Tübingen,
Morgenstelle 8,
7400 Tübingen, FRG

Received December 18, 1975; accepted April 14, 1976.

- Skolnick, P., and Daly, J. W., *Science*, **184**, 175-177 (1974).
- Segal, D. S., Kuczenski, R. J., and Mandell, A. J., *Behav. Biol.*, **7**, 75-81 (1972).
- Huang, M., Ho, A. K. S., and Daly, J. W., *Molec. Pharmacol.*, **9**, 711-717 (1973).
- Kalisher, A., Rutledge, C. O., and Perkins, J. P., *Molec. Pharmacol.*, **9**, 619-629 (1973).
- Williams, B., and Pirch, J. H., *Brain Res.*, **68**, 227-234 (1974).
- Vetulani, J., and Sulser, F., *Nature*, **257**, 495-496 (1975).
- Schultz, J., and Daly, J. W., *J. Biol. Chem.*, **248**, 843-852 (1973).
- Schultz, J., *J. Neurochem.*, **24**, 495-501 (1975).
- Gilman, A. G., *Proc. natn. Acad. Sci. U.S.A.*, **67**, 305-312 (1970).
- Hughes, I. E., Kneen, B., and Main, V. A., *J. Pharm. Pharmacol.*, **26**, 903-904 (1974).
- Burt, D. R., Enna, S. J., Creese, I., and Snyder, S. H., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 4655-4659 (1975).
- Seeman, P., Chau-Wong, M., Tedesco, J., and Wong, K., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 4376-4380 (1975).
- Farnebo, L. O., and Hamberger, B., *Acta physiol. scand. Suppl.*, **371**, 35-44 (1971).
- Stadler, H., Ciria, M. G., and Bartholini, G., *Naunyn-Schmiedeberg Arch. Pharmacol.*, **288**, 1-6 (1975).
- Gokhale, S. D., Gulati, O. D., and Parikh, H. M., *Br. J. Pharmacol.*, **23**, 508-520 (1964).

Substantia nigra of the rat contains a dopamine sensitive adenylate cyclase

A LARGE body of evidence exists to support the hypothesis that most antipsychotic drugs exert their specific actions by antagonising the effects of dopamine at postsynaptic receptors within the central nervous system¹⁻³. More recent evidence suggests that postsynaptic dopamine receptors may occur not only at sites of axon terminals, but also within the region of the cell bodies which give rise to the ascending dopaminergic projections. Ionophoretic application of dopamine to dopamine-containing neurone cell bodies in the zona compacta of the substantia nigra, inhibits firing of these cells⁴. Fluorescent histochemical examination of the substantia nigra with the glyoxylic acid technique has shown that dendrites arising from cell bodies of the zona compacta contain intensely fluorescent varicosities, and that these occur both in compacta and reticulata regions of the substantia nigra⁵. These varicosities may form the anatomical basis for a mechanism of self-inhibition whereby dopamine released from dendritic varicosities on stimulation subsequently inhibits the firing of the dopamine neurones themselves⁶. Direct support for the role of dendrites in the release of dopamine has been obtained by showing that slices of substantia nigra labelled by exposure to ³H-dopamine will release labelled dopamine when depolarised by high potassium concentrations⁷. We set out to explore further the hypothesis of self-inhibitory control by looking for biochemical evidence of a dopamine response in this region of the brain. Homogenates of tissues containing dopaminergic axon terminals respond to low concentrations of dopamine by increased production of cyclic AMP⁸⁻¹². We now provide evidence that a similar dopamine sensitive adenylate cyclase is present in the rat substantia nigra, and that it is potentially antagonised by antipsychotic drugs.

Using homogenates of tissue from substantia nigra, dissected from fresh sections of chilled tissue (Fig. 1), we found that low concentrations of dopamine were able to stimulate the production of cyclic AMP, half-maximal stimulation being obtained at approximately 2 μ M dopamine (Fig. 2). (—)Noradrenaline was about five times less potent than dopamine giving half-maximal stimulation at approximately 10 μ M. Dose-response curves for a number of related compounds were also obtained in this system, and the results are summarised in Table 1. Both the rigid dopamine analogue 2-amino-6,7-dihydroxy-(1,2,3,4)-tetrahydronaphthalene (ADTN) and epinine (*N*-methyl dopamine), gave dose-response curves similar to dopamine. These results are remarkably similar to those obtained in an

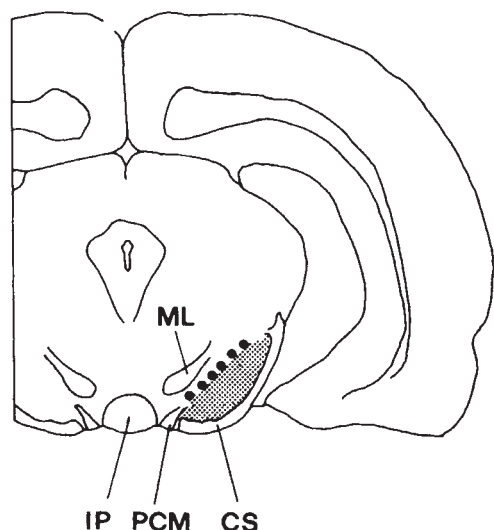


Fig. 1 Diagram of a representative section used for dissection of the substantia nigra. The coordinates for this section are A1760 μ M according to the König and Klippel atlas¹⁸. Approximate position of dopamine cell bodies of the zona compacta are shown by the black dots¹⁹. Zona reticulata is represented by stippled area. Sprague-Dawley rats weighing 180–250 g were used. After decapitation, the brains were rapidly removed and placed on ice. The brain was prepared for sectioning on a McIlwain tissue chopper in a 4 °C cold room in the following way. A horizontal cut was made through the brain removing the dorsal section of the cerebral cortex and including the dorsal portions of the basal ganglia. The brain was then placed with its cut dorsal surface face down on the stage of the tissue chopper with a piece of one-sided sticky tape on its surface, and a layer of rapid bonding glue on its upper surface. The brain was aligned to the knife and left for 10 min, its surface moistened by saline soaked filter paper. Sections 300 μ m thick were then cut through the whole substantia nigra, using the manual mechanism on the tissue chopper, and the sections transferred to ice cold microscope slides. The sections were examined under a dissecting microscope and the nigra dissected. The mean wet weight of the whole substantia nigra from both sides of the brain was 7.3 ± 0.3 mg ($n = 14$). This method was adapted from that reported by Ben-Ari and Zigmond²⁰. The tissue removed for adenylate cyclase assay was bounded by, but excluded, the medial lemniscus superiorly, pedunculus corporis mamillaris medially and the corticospinal fibres inferolaterally. ML, medial lemniscus; IP, interpeduncular nucleus; PCM, pedunculus corporis mamillaris; CS, corticospinal tract.

earlier series of experiments performed in our laboratory on the dopamine sensitive adenylate cyclase of rat striatum¹³. Furthermore, like the striatal enzyme there was a strict requirement for a catechol moiety for agonist activity, while hydroxylation at the β -position reduced activity.

Table 2 summarises the results obtained with a spectrum of antipsychotic drugs. The most potent inhibitor of the

dopamine stimulated enzyme was α -flupenthixol, followed in order of potency by chlorpromazine > haloperidol > clozapine = pimozide. The order of potency thus followed that seen for inhibition of the dopamine sensitive adenylate cyclase of the rat striatum. Comparison of the K_i values however shows that the series as a whole seems to be somewhat more potent in the substantia nigra as compared with the striatum.

When β -flupenthixol and promazine were tested at a single concentration (1 μ M) against the dopamine-stimulated adenylate cyclase, both drugs showed only weak antagonism. Promazine is known to possess only weak antipsychotic activity¹⁴ and β -flupenthixol is virtually inactive as a neuroleptic compound¹⁵.

These results provide direct biochemical evidence that dopamine may function as a neurotransmitter within the substantia nigra. Furthermore, the agonist characteristics of the dopamine sensitive adenylate cyclase in this brain region are virtually identical to those seen in the rat corpus striatum where a neurotransmitter role for dopamine is firmly established²². Our results, however, do not enable us to say whether the dopamine receptor is located on the dopamine neurone itself, or on a postsynaptic element. Thus, as well as acting on an "autoreceptor", dopamine could act by altering the firing rate of an interneurone, or by modulating the release of other neurotransmitters from presynaptic terminals located on dendrites of dopaminergic cells. Both γ -aminobutyric acid²³ and substance P (ref 24) are present in high concentrations in the substantia nigra and could be involved in such a mechanism.

Whatever the site of action of dopamine, however, the results provide direct support for the hypothesis of self inhibition in a dopaminergic nucleus and further suggest

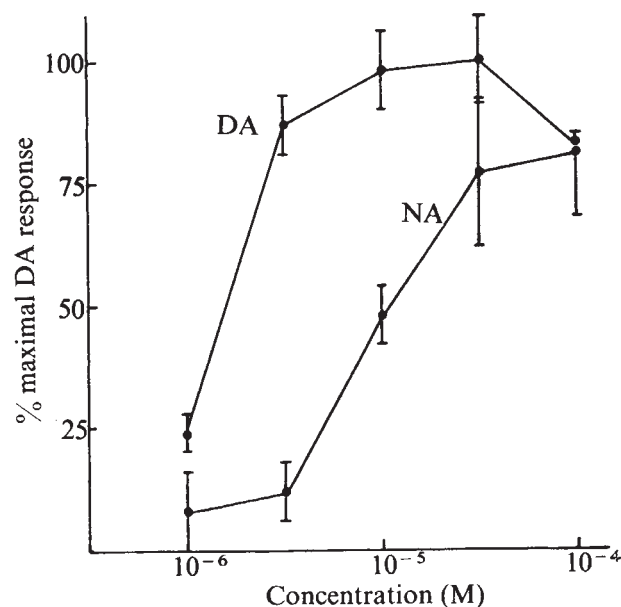


Fig. 2 Effect of dopamine (DA) and (-)-noradrenaline (NA) on stimulation of cyclic AMP production in homogenates of substantia nigra. The homogenisation and incubation procedures were essentially as described by Miller, Horn, Iversen and Pinder¹³. The procedure was, however, scaled down so that single estimations could be performed on 0.4 mg wet weight of tissue. 50- μ l aliquots of the final supernatant were assayed for cyclic AMP content by the method of Brown et al.²¹. The levels of cyclic AMP increased from a basal level of 2.82 ± 0.16 pmol per mg wet weight per min to 4.33 ± 0.27 pmol per mg wet weight per min (mean value $\pm$ s.e.m., $n = 5$) in the presence of a maximally stimulating concentration of DA (50 μ M). The experiments with NA were done by comparison with a maximally stimulating concentration of DA and the effect of 50 μ M DA normalised to 100%. Results are means $\pm$ s.e.m. of 5 separate incubations.

Table 1 Effect of dopamine and dopamine analogues on the stimulation of cyclic AMP production in homogenates of rat substantia nigra

Agonists	EC ₅₀ * (M)		Maximal stimulation (% of dopamine response)	
	S. nigra	Striatum†	S. nigra	Striatum†
Dopamine	2.3×10^{-6} M	2.0×10^{-6} M	100	100
ADTN‡	1.5×10^{-6} M	4.0×10^{-6} M	93	115
Epinephrine	2.8×10^{-6} M	1.5×10^{-6} M	107	100
(-)-Noradrenaline	1.2×10^{-5} M	4.0×10^{-5} M	81	97

($\pm$)-Isoprenaline, *p*-tyramine, (+)-amphetamine, 5-hydroxytryptamine, were all inactive at 1×10^{-4} M.

*EC₅₀ refers to the concentration required to give half maximal stimulation.

†Values quoted for striatum are taken from Miller et al.¹³.

‡2-amino-6,7,-dihydroxy-(1,2,3,4)-tetrahydronaphthalene. Basal level of cyclic AMP production was 3.24 ± 0.23 pmol per mg wet weight per min and stimulated level (5×10^{-5} M dopamine) was 4.74 ± 0.11 pmol per mg wet weight per min (mean $\pm$ s.e.m. for 4 experiments).

Table 2 Effect of neuroleptic compounds on the dopamine stimulated adenylate cyclase in rat substantia nigra homogenates

Antagonist	IC ₅₀ (M)	K _i (M)	K _i value for rat striatum (M)
α-Flupenthixol	1.1 × 10 ⁻⁸	4.8 × 10 ⁻¹⁰	1.0 × 10 ⁻⁹
Chlorpromazine	3.2 × 10 ⁻⁸	1.4 × 10 ⁻⁹	4.8 × 10 ⁻⁸
Haloperidol	1.2 × 10 ⁻⁷	5.3 × 10 ⁻⁹	9.1 × 10 ⁻⁸
Clozapine	4.0 × 10 ⁻⁷	1.7 × 10 ⁻⁸	1.7 × 10 ⁻⁷
Pimozide	4.1 × 10 ⁻⁷	1.8 × 10 ⁻⁸	1.4 × 10 ⁻⁷

Dose-response curves for the antagonists were obtained in the presence of maximally stimulating concentration of dopamine (50 μM). The neuroleptics were added immediately before the addition of dopamine. The IC₅₀ refers to the concentration of antagonist required for 50% inhibition of the increase in cyclic AMP resulting from addition of 50 μM dopamine. K_i values for each drug were calculated according to the equation $IC_{50} = K_i (1 + S/K_m)$ where *S* refers to the concentration of dopamine and *K_m* the concentration of dopamine required for half maximal stimulation of adenylate cyclase activity. Competitive inhibition has been assumed for all compounds tested (see refs 16 and 17). Samples testing for inhibition of basal activity by drug (at a concentration of 10⁻⁴ M) were run with all dose-response curves. Basal level of cyclic AMP production was 2.92 ± 0.12 pmol cyclic AMP per mg wet weight per min and stimulated level (5 × 10⁻⁵ M dopamine) was 4.59 ± 0.15 pmol cyclic AMP per mg wet weight per min (mean ± s.e.m. for 5 experiments). K_i values for rat striatum are taken from Miller, Horn and Iversen¹⁶.

that antipsychotic drugs may exert at least part of their action by interfering with this mechanism of control.

We thank Dr L. L. Iversen for reading the manuscript and Mrs J. Reed and Miss S. Gardiner for technical assistance.

O. T. PHILLIPSON
A. S. HORN

MRC Neurochemical Pharmacology Unit,
Department of Pharmacology,
Medical School, Hills Road,
Cambridge CB2 2QD, UK

Received March 1; accepted April 13, 1976.

- Matthysse, S., *Fedn Proc.*, **32**, 200-205 (1972).
- Snyder, S. H., *Archs gen. Psychiat.*, **27**, 169-179 (1972).
- Iversen, L. L., *Science*, **188**, 1084-1089 (1975).
- Aghajanian, G. K., and Bunney, B. S., *Frontiers in Catecholamine Research*, 643-648 (Pergamon, Oxford, 1973).
- Björklund, A., and Lindvall, O., *Brain Res.*, **83**, 531-537 (1975).
- Groves, P. M., Wilson, C. J., Young, S. J., and Rebec, G. V., *Science*, **190**, 522-529 (1975).
- Geffen, L. B., Jessell, T. M., Cuello, A. C., and Iversen, L. L., *Nature*, **260**, 258-260 (1976).
- Greengard, P., McAfee, D. A., and Keibadian, S. W., in *Advances in Cyclic Nucleotide Research*, **5**, 337-357 (Raven, New York, 1972).
- Brown, J. H., and Makman, M. H., *Proc. natn. Acad. Sci. U.S.A.*, **69**, 539-543 (1972).
- Keibadian, J. W., Petzold, G. L., and Greengard, P., *Proc. natn. Acad. Sci. U.S.A.*, **69**, 2145-2149 (1972).
- Horn, A. S., Cuello, A. C., and Miller, R. J., *J. Neurochem.*, **22**, 265-270 (1974).
- Von Hungen, H., and Roberts, S., *Eur. J. Biochem.*, **36**, 391-401 (1973).
- Miller, R., Horn, A., Iversen, L., and Pinder, R., *Nature*, **250**, 238-241 (1974).
- Klein, D. F., and Davis, J. M., *Diagnosis and Drug Treatment of Psychiatric Disorders* (Williams and Wilkins, Baltimore, 1969).
- Møller-Nielsen, I., et al., *Acta pharmac. tox.*, **33**, 353-362 (1973).
- Miller, R. J., Horn, A. S., and Iversen, L. L., *Mol. Pharmacol.*, **10**, 757-766 (1974).
- Clement-Cormier, Y. C., Keibadian, J. W., Petzold, G. L., and Greengard, P., *Proc. natn. Acad. Sci. U.S.A.*, **71**, 1113-1117 (1974).
- König, J. F. R., and Klippel, R. A., *The Rat Brain* (Krieger, New York, 1967).
- Ungerstedt, U., *Acta physiol. scand. Suppl.*, **367**, 1-47 (1971).
- Ben-Ari, Y., and Zigmond, R. E., *J. Physiol. Lond.*, **251**, 4P-5P (1975).
- Brown, B. L., Ekins, R. D., and Albano, J. D. M., in *Advances in Cyclic Nucleotide Research*, **2** (edit. by Greengard, P., and Robinson, G. A.), 25-40 (Raven, New York, 1972).
- York, D. H., in *Handbook of Psychopharmacology*, **6** (edit. by Iversen, L. L., Iversen, S. D., and Snyder, S. H.), 23-61 (Plenum, New York, 1975).
- Okada, Y., Nitsch-Hassler, C., Kim, J. S., Bak, I. J., and Hassler, R., *Exp Brain Res.*, **13**, 514-518 (1971).
- Duffy, M. S., Wong, J., and Powell, D., *Neuropharmacology*, **14**, 615-618 (1975).

Direct localisation of β-adrenoceptor sites in rat cerebellum by a new fluorescent analogue of propranolol

NORADRENALINE (NA) is a putative neurotransmitter, possibly involved in mediation of information transfer across synapses in the mammalian central nervous system (CNS)¹, whose postsynaptic effect is exerted, in part, by stimulation of β-adrenoceptors coupled to adenylate cyclase². Although

the rat cerebellum is known to contain NA³, and an afferent noradrenergic pathway has been found⁴⁻⁷, little information is available concerning the noradrenergic synapses there. Indirect evidence such as increase in cyclic AMP content and inhibition of spontaneous electrical activity in rat Purkinje cells⁸⁻¹⁰ (reduced by β-adrenoceptor blocking⁸), suggests, however, that such synapses are present and that the postsynaptic NA receptors are of the β type. β-adrenoceptors were identified in homogenates of whole rat cerebellum by a binding assay which used the β-adrenergic antagonist ³H-(+)-alprenolol¹¹. We have now developed a direct *in vivo* method, using 9-amino-acridin-propranolol (9-AAP), a potent fluorescent β-adrenergic blocker, as a fluorescent probe, in an attempt to detect β-adrenergic receptors in rat cerebellum.

9-AAP (D. Atlas and A. Levitzki, unpublished) is a fluorescent analogue of propranolol (*N*-[2-hydroxy-3-naphthoxy propyl]-*N'*-[9-amino acridin] isopropyl diamine) (Fig. 1). Its spectroscopic molar extinction coefficient, *E*₂₈₀, in water is 1.07 × 10⁵. The inhibitory effect of 9-AAP was

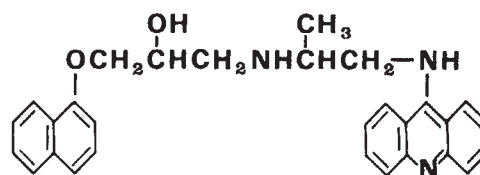


Fig. 1 The structure of 9-amino-acridin-propranolol.

calculated from the concentration of this compound which was required to inhibit 50% of the (-)-adrenaline stimulated activity in a β-receptor-dependent adenylate cyclase system. Consequently, the dissociation constant of 9-AAP was found to be (3 ± 1) × 10⁻⁸ M (Fig. 2).

9-AAP in saline (2.5 mg kg⁻¹) was administered by slow injection into the tail veins of albino rats (200-220 g). Control animals were pretreated with one of the following compounds: (+)-propranolol, (-)-propranolol, or (+)-propranolol (5 mg kg⁻¹ in saline) by slow intravenous injection. Thirty minutes later 9-AAP (2.5 mg kg⁻¹) was administered to each of the control animals. All the animals were killed by decapitation under light ether anaesthesia 30 min after injection of 9-AAP. The brains of all animals were quickly removed. The cerebellum of each animal was separated, immersed in "tissue OCT compound" and frozen in liquid nitrogen. Later, 6-8-μm coronal,

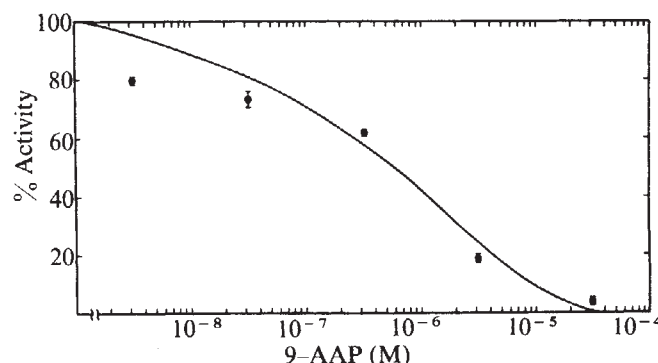


Fig. 2 Inhibition of (-)-adrenaline-stimulated activity of adenylate cyclase by 9-AAP. The dissociation constant (*K_D*) was calculated as follows^{16,17}:

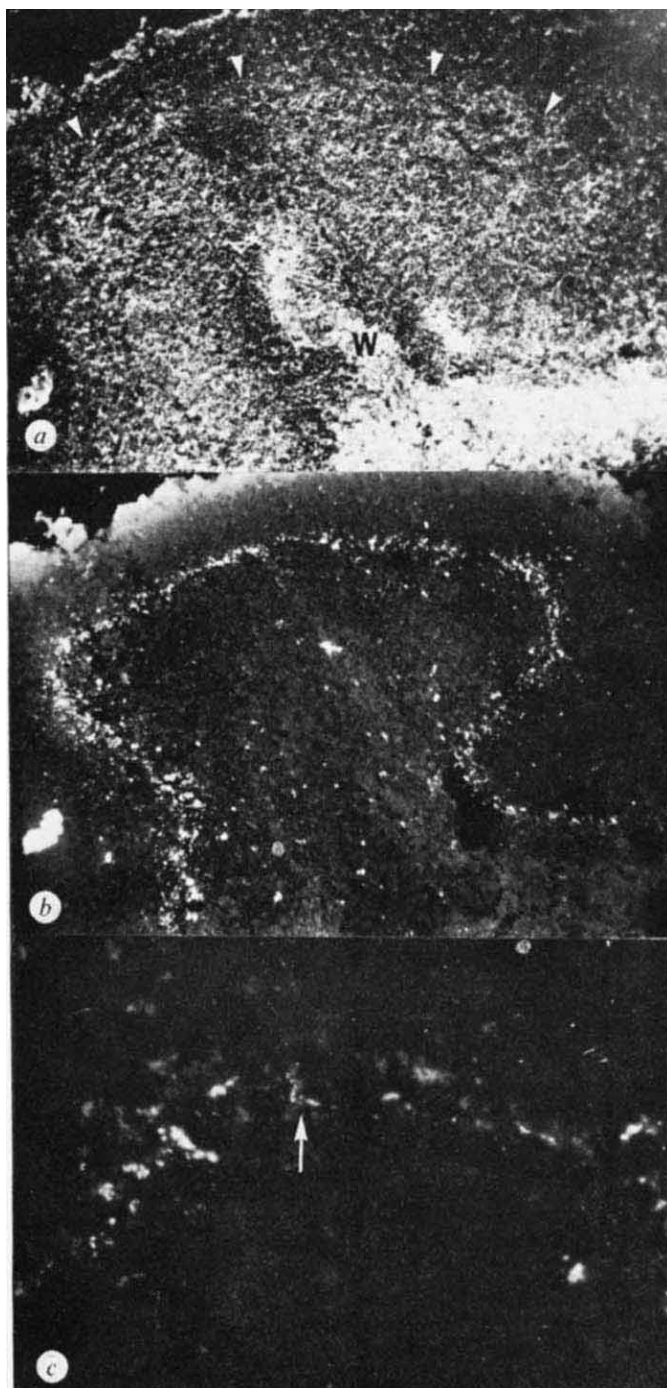
$$K_D = \frac{S_{50}}{L} \times K_L$$

where *S*₅₀ is the concentration of 9-AAP which causes 50% inhibition of total (-)-adrenaline activity; *L* is the concentration of (-)-adrenaline in the cyclase assay; *K_L* is the dissociation constant of (-)-adrenaline to the β receptor (10⁻⁶ M). *S*₅₀ = 6 × 10⁻⁷ M; *K_D* = 3 × 10⁻⁸ M.

sagittal and transverse cerebellar sections were cut in a cryostat at -20°C . The frozen sections were mounted on glass slides and air dried. Sodium phosphate buffer, 0.08 M, pH 7.4 was applied and coverslips were placed in position. The sections were visualised under phase contrast and transmitted ultraviolet illumination on a Zeiss Universal Fluorescent Microscope with HBO 200-W super pressure mercury lamp, exciter filter UG1 and barrier filter No. 41.

Dense concentration of 9-AAP binding sites was observed in every section as a well defined yellow fluorescent belt in

Fig. 3 *a*, Phase-contrast photomicrograph of sagittal section of rat cerebellum. W, white matter. The location of the Purkinje cell layer is indicated by arrows (frozen section, $\times 30$). *b*, Fluorescent photomicrograph of the same region. Intense fluorescence is observed within the Purkinje cell layer (frozen section, $\times 30$). *c*, High magnification fluorescent photomicrograph of Purkinje cell layer. Bead-like fluorescent sites are seen along apical dendrites of Purkinje cells and their bifurcations (frozen section, $\times 100$).



the cerebellar cortex. This phenomenon was localised along a line between the molecular and granular cell layers which corresponds anatomically to the Purkinje cell layer (Fig. 3*a* and *b*). Purkinje cell bodies and their major apical dendrites could be delineated by phase-contrast illumination. The highest intensity of fluorescence was observed on the apical dendrites of Purkinje cells. In most instances the distribution pattern of 9-AAP on these dendrites appeared as threads of bead-like fluorescent sites, placed at quite regular intervals from each other (Fig. 3*c*).

Less often, coalescence of these sites into larger aggregates was observed. Sparse distribution of faint fluorescent threads and dots was observed within the inner portion of the molecular cell layer. The latter finding could represent 9-AAP binding sites on fine ramifications of Purkinje cell dendrites within the molecular layer. Occasionally fluorescent spots were observed within the granular cell layer and the cerebellar white matter. Fluorescence was practically absent in cerebellar sections from control animals pretreated with ($\pm$)-propranolol and ($-$)-propranolol. In sections from control animals pretreated with (+)-propranolol, the distribution pattern of fluorescence was similar to that observed in animals which were treated by 9-AAP alone, though reduced in intensity.

It seems unlikely that the discrete anatomical localisation of the fluorescent compound in the Purkinje cell layer of rat cerebellum was due to nonspecific tissue binding. *In vitro* studies show that 9-AAP exhibits an affinity for β -adrenoceptor sites (Fig. 2). In addition, stereospecific blocking of 9-AAP binding sites was demonstrated by pretreatment with ($\pm$)- and ($-$)-propranolol but not with (+)-propranolol. The decreased intensity of fluorescence which was observed in cerebellar sections of animals pretreated with (+)-propranolol was probably due to low specific activity of the (+) isomer or to its contamination by small amounts of ($-$)-propranolol¹². Furthermore, the distribution pattern of 9-AAP binding sites correlates with the previously described localisation of the presynaptic NA-containing nerve terminals in the rat cerebellar cortex. These fibres were found in the Purkinje and molecular cell layers where they synapse with the major apical dendrites and with further dendritic processes of Purkinje cells, respectively^{6,7}. These data, in conjunction with the previous indirect evidence for the presence of β -adrenoceptors in Purkinje cells^{8-10,13-15}, suggest that the distribution of 9-AAP binding sites may correspond to β -adrenoceptor sites in rat cerebellum.

Direct binding with 9-AAP may provide a useful *in vivo* fluorescent method for the study of the postsynaptic component of NA synapses. Further studies which utilise this compound for the detection of β -adrenoceptor in other regions of the CNS and various non-neural tissues are currently in progress.

We thank Professor A. Levitzki for commenting on the manuscript. A. Itzhak and N. Conforti provided technical assistance. D.A. is supported by a fellowship from the Lady Davis Fellowship Trust and by a research grant from the Bat-Sheva de Rothschild Fund for the Advancement of Science and Technology.

ELDAD MELAMED
MOSHE LAHAV

Department of Neurology and the
Ophthalmic Pathology Laboratory,
Hadassah University Hospital and
Hebrew University Hadassah Medical School,
POB 499, Jerusalem

DAPHNE ATLAS

Department of Biological Chemistry,
Hebrew University,
20 Mamilla Road,
Jerusalem, Israel

Received March 1; accepted March 18, 1976.

- ¹ Moskowitz, M., and Wurtman, R., *New Engl. J. Med.*, **293**, 274-280 (1975).
- ² Robison, G. A., Butcher, R. W., and Sutherland, E. W., in *Cyclic AMP*, 145-231 (Academic, New York, 1971).
- ³ Anden, N. E., Fuxe, K., and Ungerstedt, U., *Experientia*, **23**, 838-839 (1967).
- ⁴ Olson, L., and Fuxe, K., *Brain Res.*, **28**, 165-171 (1971).
- ⁵ Ungerstedt, U., *Acta physiol. scand., Suppl.* 367 (1971).
- ⁶ Hokfelt, K., and Fuxe, K., *Expl Brain Res.*, **9**, 63-72 (1969).
- ⁷ Bloom, F. E., Hoffer, B. J., and Siggins, G. R., *Brain Res.*, **25**, 501-521 (1971).
- ⁸ Hoffer, B. J., Siggins, G. R., and Bloom, F. E., *Brain Res.*, **25**, 523-534 (1971).
- ⁹ Siggins, G. R., Hoffer, B. J., Oliver, A. P., and Bloom, F. E., *Nature*, **233**, 481-483 (1971).
- ¹⁰ Hoffer, B. J., Siggins, G. R., Oliver, A. P., and Bloom, F. E., *J. Pharmac. exp. Ther.*, **184**, 553-569 (1973).
- ¹¹ Alexander, R. W., Davies, J. N., and Lefkowitz, R. J., *Nature*, **258**, 437-440 (1975).
- ¹² Howe, R., and Shanks, R. G., *Nature*, **210**, 1336-1338 (1966).
- ¹³ Siggins, G. R., Battenberg, E. F., Hoffer, B. J., and Bloom, F. E., *Science*, **179**, 585-588 (1973).
- ¹⁴ Siggins, G. R., Hoffer, B. J., and Bloom, F. E., *Science*, **165**, 1018-1020 (1969).
- ¹⁵ Siggins, G. R., Hoffer, B. J., and Bloom, F. E., *Brain Res.*, **25**, 535-553 (1971).
- ¹⁶ Atlas, D., Steer, M. L., and Levitzki, A., *Proc. natn. Acad. Sci. U.S.A.*, **71**, 4246-4248 (1974).
- ¹⁷ Levitzki, A., Sevilla, N., Atlas, D., and Steer, M. L., *J. molec. Biol.*, **97**, 35-53 (1975).

Spontaneous electrical potentials and pituitary hormone (MSH) secretion

CELLULAR secretions from either exocrine or endocrine glands are regulated in most cases by hormonal or neuro-hormonal agents. These chemical messengers interact with cellular receptors to stimulate a particular secretory response. In general, it would seem that changes in the transmembrane electrical potential of the secretory unit is in some way intimately involved in this stimulus-secretion coupling¹. Here we provide evidence of cellular secretion resulting from spontaneous electrical activity (most probably depolarisations) after removal of a previously inhibitory stimulus.

Melanophore stimulating hormone is synthesised and released from the pars intermedia of the vertebrate pituitary². This hormone has an important role in integumental colour changes of most vertebrate species³. MSH secretion from the vertebrate pituitary is regulated, as for other pituitary hormones, by hypothalamic control². Transplantation of the pituitary gland to an ectopic site or to an *in vitro* condition results in a continuous uninhibited release of MSH⁴ and prolactin⁵. The *in vitro* secretion of these hormones is reversibly inhibited by a variety of experimental methods demonstrating that such manipulations are not themselves responsible for this enhanced hormone secretion^{5,6}. Thus, MSH and prolactin secretion are unique in that isolation from a hypothalamic inhibitory influence results in active, spontaneous secretion of these hormones.



Fig. 1 Continuous oscilloscope recording of isolated neuro-intermediate lobe from *R. pipiens* pituitary. The top sweep indicates extracellular spontaneous electrical activity from a single unit in the pars intermedia. At the time indicated by the arrow, a 0.8 V, 10 ms pulse from an electronic stimulator was passed through the gland by way of two silver-silver chloride electrodes flanking the pars intermedia. The bottom trace shows the increase in frequency of spikes occurring approximately two seconds later with twice as much activity as the spontaneous sweep and lasting for about 30 s. This enhancement of electrical activity is correlated with an increase in hormone secretion (Fig. 3). The vertical amplitude scale is 25 mV and the horizontal time scale is 150 ms.

Other cell types require exogenous stimulatory cues to enhance secretion.

Although several models for the neuronal control of pars intermedia cell function have been provided^{2,7}, the physiological basis for the autonomous secretion of MSH in the absence of hypothalamic control is unresolved. As was earlier suggested for the prolactin cell of the pars distalis⁸, MSH secretion may result from a spontaneous depolarisation of pars intermedia cells. Evidence that such a hypothesis may be correct is presented here.

Neurointermediate lobes of the frog, *Rana berlandieri forreri* (*Rana pipiens, sensu lato*) were removed and incubated in tissue culture for one to two weeks. The neuro-intermediate lobes were cultured to be reasonably certain that when recording electrical activity of the pars intermedia cells, injury potentials set up by cut axonal endings would not be present and be confused with the electrical activity of the pars intermedia cells. Interneurons are apparently absent from the pars intermedia⁹. To further insure against such a possibility, the intermediate lobes in some experiments were incubated in the presence of tetrodotoxin (TTX) which is known to inhibit action potentials in nerves. Tetrodotoxin (10^{-6} M) did not block electrical activity or inhibit MSH secretion.

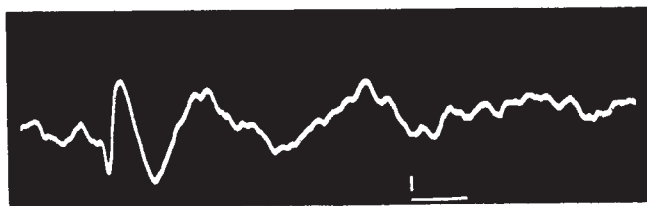


Fig. 2 Fast oscilloscope sweep showing the shape of a single spike unit as in Fig. 1. The triphasic (positive-negative-positive) wave is characteristic of a depolarisation recorded extracellularly from the pars intermedia. Triphasic pulses are generally associated with conduction of excitation along nerve axons. Similar pulses recorded in these experiments are difficult to explain except perhaps that potentials travelling through cells in the pars intermedia may be electrically coupled. As a control, no recordings could be found in the pars nervosa cultured *in vitro*. The vertical amplitude scale is 25 mV and the horizontal time scale is 10 ms.

Spontaneous electrical activity (Fig. 1, top) was recorded with 3 M KCl-filled glass micropipette electrodes (0.5-1.5 μ m diameter, 10-30 M Ω resistance) from just inside the periphery of the gland down to about 300 μ m. A silver reference electrode was placed in the bathing amphibian Ringer's solution. Although quite often the tip of the electrode would penetrate a cell with a resulting rapid negative deflection, no steady intracellular potentials could be recorded for any length of time. Instead, extracellular spikes (as shown in Fig. 2) indicate characteristically depolarising potentials with triphasic events (positive-negative-positive, with major phase mostly negative). Whenever a fairly steady frequency of spikes was found, a 0.8 V, 10 ms single pulse of electricity was sent through the gland by way of two silver-silver chloride stimulating electrodes flanking opposite sides of the neurointermediate lobe. In about 75% of the more than 90 cases studied, the electrical stimulation tended to increase the frequency of spikes two- to threefold for as long as 60 s (Fig. 1, bottom) while the remaining 25% maintained the same frequency or showed inhibited frequency. The latency between stimulation and onset of frequency changes varied between a half to a few seconds and very possibly the range might extend further in either direction. The latency between a single electrical stimulus and the increase in frequency of spontaneous pulses as well as the duration of the after-discharges are much greater than would be expected if the pulses were of neuronal origin. Thus, in the majority of cases, spontaneous

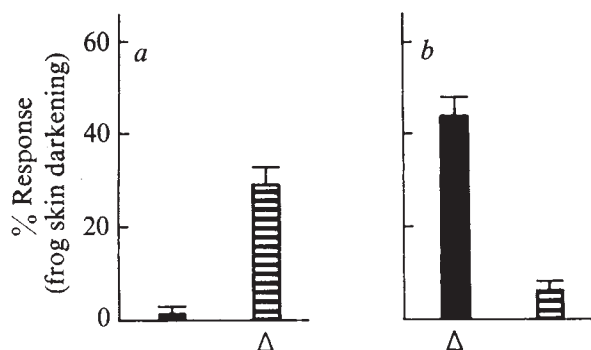


Fig. 3 Bioassay of MSH secreted from isolated, tissue-cultured frog neurointermediate lobes. Eight glands were used and were divided into two experimental groups. After an initial 30-min preincubation (to equalise secretion rates in both groups), one group of glands (arrow, *a*) was subjected to 0.8 V, 10 ms pulses at 1 pulse s^{-1} for 60 min. The other group of glands remained as a control without stimulation. Then, in a subsequent incubation, a similar stimulation was applied to the previously untouched group of glands (arrow, *b*) and the other previously stimulated glands served as controls. Note the correlation between electrical stimulation and enhancement of MSH secretion and reversal of secretion when stimulation is terminated. Each value represents the mean $\pm$ s.e., darkening response of the frog skins (8 per group) used to bioassay the released MSH¹⁵.

electrical activity of pars intermedia cells is enhanced through electrical stimulation.

Figure 3 shows the relative quantities of MSH released during electrical stimulation of tissue-cultured neurointermediate lobes. These results indicate that electrical stimulation reversibly enhances MSH secretion. The electrophysiological readings and the bioassay of released MSH were repeated several times with the same conclusion using both tissue-cultured glands and freshly excised glands. Similar data (electrically enhanced pulse frequency and enhanced MSH secretion) in our preliminary experiments using the isolated rat pituitary gland have been obtained.

These results suggest that MSH secretion from the isolated frog or rat pituitary may result from spontaneous depolarisations of pars intermedia cells. Since Ca^{2+} is required for MSH secretion², depolarisation of the cell membrane may act to increase the permeability to this ion and activate the secretion process. Evidence that depolarisation itself leads to enhanced MSH secretion is shown by the demonstration that electrical stimulation of isolated neurointermediate lobes results in reversible enhancement of hormone secretion. The results demonstrate a unique mechanism of pars intermedia cell secretion possibly only shared by a similar mechanism for prolactin release in the pars distalis. Whereas enhanced cellular secretion above basal levels from other cellular species is induced by changes in membrane potential resulting from hormonal, neurohormonal or pharmacological stimulation, secretion of MSH results spontaneously when the pars intermedia cells escape from hypothalamic regulatory cues and depolarise.

The nature of the hypothalamic control of MSH secretion is at present debatable. Both peptide factors of hypothalamic origin and direct innervation of the pars intermedia by hypothalamic neurones are considered to regulate MSH secretion. There is clear morphological evidence for a control of pars intermedia cell function by direct innervation⁹. Light⁹ and electron microscopic evidence^{10,11}, favours the view that the inhibitory effect of these neurones is mediated by a catecholamine. The physiological support for such morphological evidence was provided by the demonstration that catecholamines such as noradrenaline, adrenaline and dopamine inhibit MSH secretion reversibly *in vitro*^{12,13}. The receptor basis for the mechanism of action of these catecholamines has also been provided¹³.

Our results provide evidence that inhibition of MSH secretion is maintained by a tonic polarisation of the individual pars intermedia cells. Photostimulatory cues of ocular and possibly epithalamic (possibly pineal) origin modulate or otherwise influence the hypothalamic input to the pars intermedia^{8,14}. It has been suggested that each pars intermedia cell may be innervated by an inhibitory neurone. Inhibition of these neurones might be accomplished by other neurones, possibly directly or by presynaptic relationships. Variable activity of the latter neurone would modulate the activity of the tonically inhibitory neurone^{11,14}. In the absence of the inhibitory polarising influence on the pars intermedia cells, they would spontaneously depolarise and this would result in the release of MSH. Although a number of secretory systems have been studied¹ (pars distalis of pituitary, salivary glands, exocrine pancreas), we know of no other example where cellular secretions may be initiated by such spontaneous depolarisations. The pars intermedia cell, therefore, offers a unique model system for studying various electrophysiological parameters.

We thank Dr Ted Glatke and Dr Peter Pickens for their help. This study was supported, in part, by a grant from the US Public Health Service.

M. DUFF DAVIS
MAC E. HADLEY

Department of Cell and Developmental Biology,
University of Arizona,
Tucson, Arizona 85721

Received December 8, 1975; accepted April 5, 1976.

- 1 Matthews, E. K., in *Secretory Mechanisms of Exocrine Glands* (edit. by Thorn, N. A., and Paterson, O. H.), 185-198 (Academic Press, New York, 1974).
- 2 Hadley, M. E., and Bagnara, J. T., *Am. Zool.*, **15**, 81-104 (1975).
- 3 Bagnara, J. T., and Hadley, M. E., *Chromatophores and Color Change: the Comparative Physiology of Animal Pigmentation* (Prentice-Hall, Englewood Cliffs, New Jersey, 1973).
- 4 Etkin, W., *Gen. comp. Endocr. Suppl.*, **1**, 148-159 (1962).
- 5 Parsons, J. A., *J. Physiol., Lond.*, **210**, 973-987 (1970).
- 6 Hadley, M. E., Bower, Sr. A., and Hruby, V. J., *Yale J. Biol. Med.*, **46**, 602-616 (1973).
- 7 Dierst-Davies, K., Ralph, C. L., and Pechersky, J. L., *Gen. comp. Endocr.*, **6**, 409-419 (1966).
- 8 Dawson, D. C., and Ralph, C. L., *Neuroendocrinology*, **15**, 267-280 (1974).
- 9 Enemar, A., and Falck, B., *Gen. comp. Endocr.*, **5**, 577-583 (1965).
- 10 Hopkins, C. R., *Gen. comp. Endocr.*, **16**, 112-120 (1971).
- 11 Nakai, Y., and Gorbman, A., *Gen. comp. Endocr.*, **13**, 108-116 (1969).
- 12 Iturriza, F. C., and Kasal-Iturriza, M., *Gen. comp. Endocr. Suppl.*, **3**, 108-113 (1972).
- 13 Bower, A., Hadley, M. E., and Hruby, V. J., *Science*, **184**, 70-72 (1974).
- 14 Oshima, K., and Gorbman, A., *Gen. comp. Endocr.*, **13**, 98-107 (1969).
- 15 Huntington, T., and Hadley, M. E., *Endocrinology*, **96**, 472-479 (1974).

Evidence for analgesic activity of enkephalin in the mouse

HUGHES *et al.*¹ have recently described the identification and structure elucidation of methionine- and leucine-enkephalin, the postulated endogenous morphine-like factors from mammalian brain²⁻⁴. The structures were confirmed by comparing the mass spectra and electrophoretic mobilities of natural and synthetic material and their biological activities established *in vitro* in preparations of the mouse vas deferens and the guinea pig ileum. We have synthesised the two pentapeptides H-Tyr-Gly-Gly-Phe-Met-OH and H-Tyr-Gly-Gly-Phe-Leu-OH, as well as two analogues (desamino-Tyr)-Gly-Gly-Phe-Met-OH and H-Tyr-Gly-Gly-OH and compared their biological activities. Our results indicate that the two natural compounds induce analgesia in mice, whereas the two analogues are devoid of such activity.

The pentapeptides were built up by fragment condensation methods⁵ (mixed anhydride or *N,N'*-dicyclohexylcarbodiimide and 1-hydroxybenzotriazole). Boc-Tyr-OH was coupled with H-Gly-Gly-OBz to give Boc-Tyr-Gly-Gly-OBz (melting point 97-102 °C, $[\alpha]_D^{20} = -5.4^\circ$, $c = 2$ in DMF). After catalytic hydrogenation, Boc-Tyr-Gly-Gly-OH was reacted either with H-Phe-Met-OMe or H-Phe-Leu-OEt to produce Boc-Tyr-Gly-Gly-Phe-Met-OMe (melting point 92 °C, $[\alpha]_D^{20}$

= -20.2°, $c = 2$ in DMF) and Boc-Tyr-Gly-Gly-Phe-Leu-OEt (melting point 95 °C, $[\alpha]_D^{20} = -16.1^\circ$, $c = 2$ in DMF). After saponification of the esters, the Boc- groups were cleaved by trifluoroacetic acid. The two natural enkephalins were obtained as homogenous crystalline compounds after transforming the peptides to a salt free form on an ion exchange column (methionine-enkephalin, melting point 190–196 °C, $[\alpha]_D^{20} = +26.1^\circ$, $c = 1$, 95% AcOH; leucine-enkephalin, melting point 157–167 °C, $[\alpha]_D^{20} = +26.4^\circ$, $c = 1$, 95% AcOH). The desamino-tyrosine analogue of leucine-enkephalin was synthesised in a similar manner starting from 3-(4-hydroxyphenyl) propionic acid. All peptides gave correct amino acid analysis.

The analgesic effects of methionine- and leucine-enkephalin were studied using the tail flick test in mice⁸. A direct comparison with morphine using the same pretreatment interval was complicated by their different pharmacokinetic properties. Morphine (0.56–3.2 µg per mouse) had a slow onset of action following intracerebroventricular administration, peak activity being reached after 15 min. In contrast the two endogenous enkephalins were maximally active after 2 min and had a duration of action of barely 5 min. This difference was not unexpected since enkephalin is readily destroyed enzymatically^{2,4}. Doses of morphine that had an analgesic effect within 2 min were toxic at later observation times.

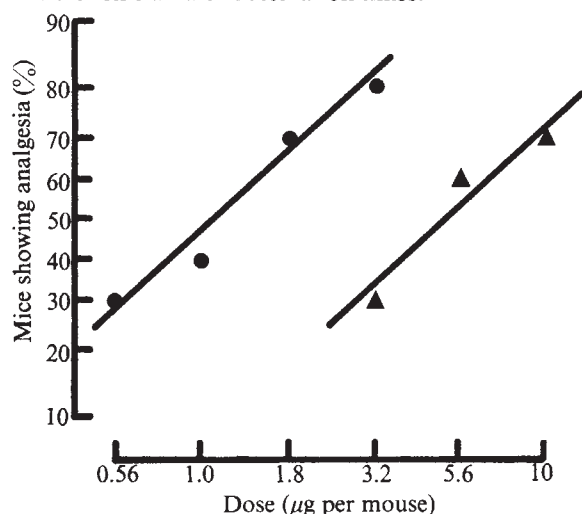


Fig. 1 The analgesic effect of morphine in the tail flick test⁸ in the mouse 15 min after intracerebroventricular administration alone (●) and following pretreatment with 0.1 mg kg⁻¹ subcutaneous naloxone (▲). 10 animals per dose.

The dose-response curves of the analgesic effect of morphine at 15 min and of the two enkephalins at 2 min were approximately parallel (Figs 1 and 2). Morphine was 75 and 240 times more active than methionine- and leucine-enkephalin respectively (Table 1). Naloxone (0.1 mg kg⁻¹ subcutaneously), administered 20 min before the measurement of analgesia, inhibited the analgesic effects of methionine-enkephalin and morphine in a competitive manner, the curves being displaced to the right by a factor of about 5 (Figs 1 and 2a). This suggested that the

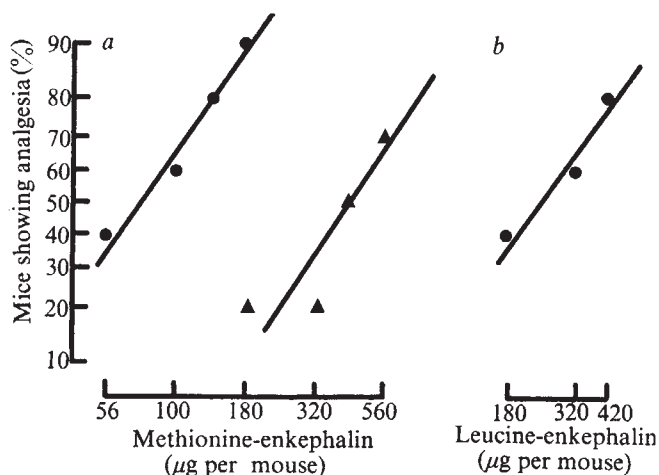


Fig. 2 The analgesic effects of methionine- and leucine-enkephalin in the tail flick test⁸ 2 min after intracerebroventricular administration alone (●) and following pretreatment with 0.1 mg kg⁻¹ subcutaneous naloxone (▲). 10 animals per dose.

two compounds were affecting the same system in a similar manner. In contrast, the analgesic effect of leucine-enkephalin was not inhibited by pretreatment with naloxone except at the very high dose of 10 mg kg⁻¹ subcutaneously. The different susceptibilities of the two enkephalins to blockade by naloxone could indicate that the leucine analogue may behave as a mixed agonist-antagonist at the opiate receptor⁹, whereas the methionine form may more closely resemble morphine in its action. The results may explain why Hughes *et al.*¹⁰ required significantly more naloxone to antagonise the effects of the unseparated enkephalin than to inhibit the effect of normorphine *in vitro*.

The maximum effect of morphine (1.0–3.2 mg kg⁻¹), administered intravenously, occurred after 15 min (ED₅₀ = 2.0 mg kg⁻¹). Methionine-enkephalin exhibited a transient effect at very high doses (ED₅₀ = 170 mg kg⁻¹ intravenously), analgesic activity being detected only within the first 15 s of administration. Leucine-enkephalin was only weakly active (4 out of 10 mice showing analgesia) at the highest soluble dose (320 mg kg⁻¹ intravenously). These results obtained *in vivo* with the two enkephalins following intracerebroventricular and intravenous administration support the *in vitro* finding¹ that the methionine analogue is more active than the leucine form.

In binding studies using rat brain homogenate¹¹, methionine-enkephalin at 25 °C was twelve times less potent than morphine in inhibiting specific ³H-naloxone binding in a sodium-free medium and approximately six times less potent in a medium containing sodium (Table 2). Hughes *et al.*¹ claimed a higher affinity for enkephalin than morphine but they used guinea pig and not rat brain homogenate. Furthermore, the temperature at which incubation and filtration is carried out has a pronounced effect on the binding of opiates¹² which might be responsible for the discrepancy between the results.

Table 1 Effect of naloxone pretreatment on analgesic activity of morphine and enkephalin

Test substance	Time (min) after administration	ED ₅₀ of analgesic activity µg per mouse intracerebroventricularly ($P = 0.05$)	
		Alone	After naloxone
Morphine-HCl	15	1.1 (0.7–1.8)	5.2 (3.2–8.3)
H-Tyr-Gly-Gly-Phe-Met-OH	2	75 (52–100)	420 (305–579)
H-Tyr-Gly-Gly-Phe-Leu-OH	2	240 (159–363)	180 (115–282)
(des-NH ₂ -Tyr)-Gly-Gly-Phe-Leu-OH	2	> 32 (toxic)	—
H-Tyr-Gly-Gly-OH	2	≥ 320	—

ED₅₀ values ($P = 0.05$) for the analgesic activity in the tail flick test in the mouse alone and after pretreatment with naloxone (0.1 mg kg⁻¹ subcutaneously). The ED₅₀ is defined as the dose that prolonged the reaction time in seconds by more than 75% in half of the mice. 10 animals were used per dose and drugs administered intracerebroventricularly (10–30 µl per mouse) were dissolved in artificial cerebrospinal fluid⁷. The accuracy of the intracerebroventricular injections, performed using the method of Haley and McCormick⁸, were confirmed macroscopically with the aid of Indian ink.

Table 2 Binding of ^3H -naloxone by rat brain

Test substance	IC ₅₀ of specific ^3H -naloxone binding No NaCl nM	100 mM NaCl nM	IC ₅₀ ratio + NaCl/- NaCl
Morphine-HCl	4.5 ± 0.6 (3)	73 ± 10 (3)	16.2
Pentazocine	12.3 ± 0.3 (2)	87 ± 3 (2)	7.1
H-Tyr-Gly-Gly-Phe-Met-OH	55.4 ± 9.4 (5)	412 ± 69 (5)	7.4
H-Tyr-Gly-Gly-Phe-Leu-OH	224 ± 19 (5)	1,075 ± 75 (5)	4.8
(des NH ₂ -Tyr)-Gly-Gly-Phe-Leu-OH	> 10,000	> 10,000	—
H-Tyr-Gly-Gly-OH	> 10,000	> 10,000	—

IC₅₀ values for the inhibition of specific ^3H -naloxone binding in rat brain homogenate. The IC₅₀ is defined as the concentration of substance required to inhibit by 50% specific binding of ^3H -naloxone (nM) to preincubated (30 min at 37 °C) homogenates of rat brain minus cerebellum in 0.05 M Tris buffer (pH 7.4 at 25 °C)¹². Incubations in triplicate were assayed at 25 °C for 40 min alone or in the presence of 100 mM NaCl. Filtrations were performed at 25 °C. The values are the mean ± s.e.m.; the number of experiments is given in parentheses. The "sodium response ratio" is defined as the ratio of the IC₅₀ values in the presence and absence of sodium¹¹.

The low "sodium response ratio" calculated for leucine-enkephalin is characteristic for a dual agonist-antagonist at opiate receptors¹¹ and corresponds with the findings *in vivo*. On the other hand, the similarity between the sodium response ratio of methionine-enkephalin and the partial agonist, pentazocine, was surprising and suggests that methionine-enkephalin might also possess antagonistic activity. Further studies are in progress to examine in greater detail the opiate receptor blocking activities of the two enkephalins and their stabilities during incubation.

The desamino analogue in non-toxic doses and the tripeptide did not produce analgesia following intracerebroventricular administration (Table 1) and had no affinity for the opiate receptor in the binding studies (Table 2). The results with the former compound show the terminal amino group of enkephalin to be indispensable for its morphine-like activity and suggest that it corresponds to the nitrogen atom of morphine. The inactivity of the tripeptide shows that the minimal requirements for analgesic activity are not met by the first three amino acids of enkephalin.

HEINZ H. BÜSCHER
RONALD C. HILL
DIETMAR RÖMER*

Biological and Medical Research Division,

FRANCIS CARDINAUX
ANNEMARIE CLOSSE
DANIEL HAUSER
JANOS PLESS

Chemical Research Pharmaceutical Division,
Sandoz Ltd,
CH-4002 Basle,
Switzerland

Received March 3; accepted April 13, 1976.

*To whom reprint requests should be addressed.

¹ Hughes, J., *et al.*, *Nature*, **258**, 577-579 (1975).

² Hughes, J., *Brain Res.*, **88**, 295-308 (1975).

³ Terenius, L., and Wahlström, A., *Acta physiol. scand.*, **94**, 74-81 (1975).

⁴ Pasternak, G. W., Goodman, R., and Snyder, S. H., *Life Sci.*, **16**, 1765-1769 (1975).

⁵ Houben-Weyl, in *Methoden der Organischen Chemie*, **15** (Georg Thieme, Stuttgart, 1974).

⁶ D'Amour, D. E., and Smith, D. L., *J. Pharmac. exp. Ther.*, **72**, 74-78 (1941).

⁷ Palaic, D., Page, I. H., and Khairallah, P. A., *J. Neurochem.*, **14**, 63-69 (1967).

⁸ Haley, T. J., and McCormick, W. G., *Br. J. Pharmac.*, **12**, 12-15 (1957).

⁹ Kosterlitz, H. W., Lord, J. A. H., and Watt, A. J., in *Agonist and Antagonist Actions of Narcotic Analgesic Drugs* (edit. by Kosterlitz, H. W., Collier, H. O. J., and Villarreal, J. E.), 45-61 (Macmillan, London, 1972).

¹⁰ Hughes, J., Smith, T., Morgan, B., and Fothergill, L., *Life Sci.*, **16**, 1753-1758 (1975).

¹¹ Pert, C. B., and Snyder, S. H., *Molec. Pharmac.*, **10**, 868-879 (1974).

¹² Creese, I., Pasternak, G. W., Pert, C. B., and Snyder, S. H., *Life Sci.*, **16**, 1837-1842 (1975).

agonist activity peripherally⁷, raises the question of the possible physiological significance of this substance in the central nervous system. There is evidence from a number of studies suggesting that the brainstem is an important site for antinociceptive action³⁻⁵. In previous studies⁶, morphine applied iontophoretically to single neurones in the rat brainstem, has been found to cause either an increase or a decrease in firing rate, but only the depressant response showed stereospecificity and could be blocked by the specific opiate antagonist, naloxone^{7,8}. Similar findings on cortical neurones have been reported⁹. We have therefore investigated the effects of iontophoretically applied methionine-enkephalin (met-enkephalin), the more potent peripherally of the two pentapeptides, on the activity of single neurones in the rat brain stem and compared these effects with those of morphine, and also of another narcotic analgesic, etorphine, which is considerably more potent. Since naloxone has been shown to block peripheral effects of enkephalin¹⁰, we have examined its effects on the neuronal responses to met-enkephalin, morphine and etorphine. Our findings support the concept of a physiological role for endogenous enkephalin in the central nervous system.

Table 1 Comparison of effects of enkephalin with either morphine or etorphine applied to same neurone

	Morphine				Etorphine	
	Effect +	+	NE	-	+	-
	9	0	0	0	0	0
Enkephalin (NE)	0	7	0	0	2	0
-	14	10	10	3	0	8
-+	1	0	1	0	0	0

Numbers indicate number of neurones in each case.

+, Excitation; -, inhibition; NE, no effect; + - excitation followed by inhibition; - +, inhibition followed by excitation.

Drugs were applied from five-barrelled glass micropipettes inserted into the medullary reticular formation of rats under light urethane anaesthesia (see ref. 6 for further experimental details). Met-enkephalin (5 mM, pH 5.0) supplied by Dr B. A. Morgan of Reckitt & Colman, and etorphine HCl (24 mM, pH 5.0) were dissolved in distilled water. Other drug solutions were as described previously⁷.

The effects of met-enkephalin, applied iontophoretically, for periods of 5-90 s with positive currents of 10-45 nA were studied on 82 spontaneously active neurones in 21 rats. Enkephalin depressed the activity of 70% (58) of the neurones tested (Fig. 1a) and excited 11% (9). Biphasic effects were occasionally seen (4) but whereas the biphasic response to morphine consisted of excitation followed by inhibition, that to enkephalin was consistently inhibition followed by excitation. In Table 1 the effects of enkephalin are compared with those of either morphine or etorphine, applied to the same neurone. When morphine or etorphine produced depression, so did enkephalin. On the other hand, although enkephalin mimicked some of the excitatory responses to morphine, in many cases it depressed neurones which were excited by morphine. There were also some

Effects of microiontophoretically applied methionine-enkephalin on single neurones in rat brainstem

THE identification of a specific opiate receptor in mammalian brain¹, and also of an endogenous morphine-like factor, enkephalin, which possesses potent opiate

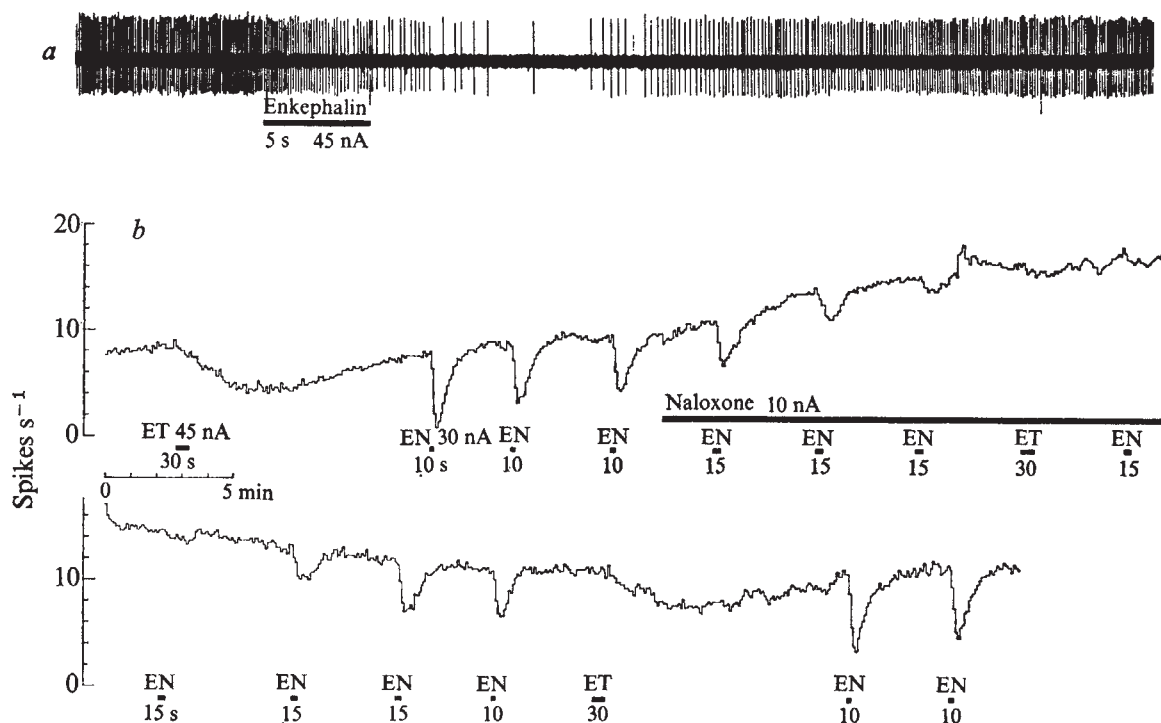


Fig. 1a, Inhibitory effect of iontophoretically applied enkephalin (45 nA for 5 s; current balancing was used) on the spontaneous activity of a rat medullary neurone. The inhibitions caused by enkephalin were usually slow in onset, reaching a maximum effect 10–40 s after the start of application, and recovering slowly over 1–3 min from the end of application. **b**, Antagonism by naloxone of the inhibitory actions of etorphine (ET) and enkephalin (EN) on the firing rate (spikes s⁻¹ over consecutive 5-s epochs) of a rat medullary neurone. Periods of iontophoretic applications are indicated by the bars; expelling currents (balanced; nA) are shown above, and durations of applications below the bars. Note that during and after naloxone application, longer enkephalin applications were used, but in spite of this, the effects of enkephalin and of etorphine were blocked; gradual and complete recovery was observed.

neurones, depressed by enkephalin, which seemed to be unaffected by application of morphine. Unlike morphine, etorphine was never observed to increase neuronal activity, and therefore the correlation between the effects of enkephalin and those of etorphine was much better.

Naloxone was applied iontophoretically, using currents of 5–30 nA and blocked enkephalin-induced depression 8 out of 8 times (Fig. 1b). In two of these, morphine depression was also blocked by naloxone and in another two etorphine depression was blocked. In a single neurone which was excited by both morphine and enkephalin, naloxone failed to block either response. Naloxone sometimes had effects on firing rate but these were unrelated to its antagonistic actions.

It would seem that in the brainstem the depressant action of enkephalin is likely to be more important than its excitatory action since this response was reversed by naloxone. This is in good agreement with the previous finding that only the depressant action of morphine was antagonised by naloxone. Note that etorphine, a more potent analgesic drug, produced only depression of neuronal activity and this too was antagonised by naloxone. Whether or not the depressant response to enkephalin can be taken to indicate a physiological action of this substance on these neurones, remains to be seen. Until more information is available on the distribution of enkephalin at the subcellular level, this will be an open question.

L.A.L. is an MRC scholar.

P. B. BRADLEY
I. BRIGGS
R. J. GAYTON
LYNN A. LAMBERT

Department of Pharmacology
and MRC Neuropharmacology Unit,
Medical School,
Birmingham B15 2TJ, UK

Received March 16; accepted April 23, 1976.

- ¹ Snyder, S. H., *Nature*, **257**, 185–189 (1975).
- ² Hughes, J., *et al.*, *Nature*, **258**, 577–579 (1975).
- ³ Benjamin, R. M., *Brain Res.*, **24**, 525–529 (1970).
- ⁴ Herz, A., Albus, K., Metys, J., Schubert, P., and Teschemacher, H., *Neuropharmacology*, **9**, 539–551 (1970).
- ⁵ Mayer, D. J., Wolfe, T. L., Akil, H., Carder, B., and Liebeskind, J. C., *Science*, **174**, 1351–1354 (1971).
- ⁶ Bradley, P. B., and Dray, A., *Br. J. Pharmacol.*, **50**, 47–55 (1974).
- ⁷ Bramwell, G. J., and Bradley, P. B., *Brain Res.*, **73**, 167–170 (1974).
- ⁸ Bradley, P. B., and Bramwell, G. J., *Br. J. Pharmacol.*, **53**, 462P (1975).
- ⁹ Satoh, M., Zieglansberger, W., Fries, W., and Herz, A., *Brain Res.*, **82**, 378–382 (1974).
- ¹⁰ Hughes, J., Smith, T., Morgan, B., and Fothergill, Linda, *Life Sci.*, **16**, 1753–1758 (1975).

Effects of methionine-enkephalin and leucine-enkephalin compared with those of morphine on brainstem neurones in cat

Two new peptides, methionine-enkephalin (met-enkephalin) and leucine-enkephalin (leu-enkephalin), have been isolated from brain, identified and synthesised¹. The interest in these compounds lies in their morphine-like activity which was recognised and tested by their action on isolated preparations from the peripheral nervous system. Since morphine is used for its central action in producing analgesia, it is important to investigate whether these substances have activity in the central nervous system similar to that of morphine. We report here the findings of similar depressant actions of all three substances on brainstem neurones in the cat.

To compare the actions of the enkephalins and morphine we applied them to single neurones in the brainstem of unanaesthetised, decerebrate cats by iontophoresis from five- or seven-barrelled micropipettes. Pipettes were inserted through the floor of the IVth ventricle 2–7 mm rostral to the obex and 0–1.5 mm lateral from midline. Histological sections showed that the pipettes sampled

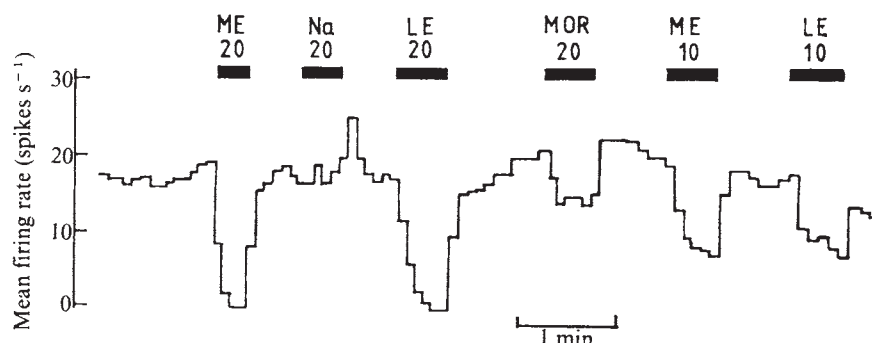
Table 1 Percentages of brainstem neurones showing inhibition, a biphasic effect of inhibition followed by excitation, or no effect, in response to the iontophoretic application of enkephalin or morphine

	Inhibition	Biphasic	No effect
Met-enkephalin ($n = 96$)	89	1	10
Leu-enkephalin ($n = 40$)	92.5	0	7.5
Morphine ($n = 84$)	87	2	11

mainly the activity of neurones in nucleus reticularis gigantocellularis and pontis caudalis. The substances, which were usually applied with currents of 10–30 nA for 30-s periods, were tested for effects on spontaneous activity or, when this was low or absent, on activity induced by the continuous application of homocysteic acid (0–17 nA).

The predominant effect of met- and leu-enkephalin and morphine was depression of neuronal firing rate as shown in Fig. 1. Rarely, the depression was followed by excitation, but we have not observed any instances of excitation alone. A few units did not show any response, and the incidence of the different effects is given in Table 1. In 81 out of 84 cells in which the effects of met-enkephalin and morphine were compared, their effects were qualitatively the same (73 inhibited, in one of which the inhibition was followed by excitation, and 8 unaffected). The effects of met- and leu-enkephalin began within the first 5 s of the ejection period and usually reached a peak by 15–20 s.

Fig. 1 Effects of iontophoretically applied met-enkephalin (ME), leu-enkephalin (LE) and morphine (MOR) on the firing rate, induced by a continuous current of 9 nA of DL-homocysteic acid, of a brainstem neurone. The application of Na^+ served as a control for the effects of current. Applications are indicated by the horizontal bars. Currents are shown in nA. Firing rate shown in successive 5-s epochs. Met- and leu-enkephalin were supplied by Dr B. A. Morgan of Reckitt & Colman; they were dissolved at pH 4.5, at 0.016 M. Morphine was used as the sulphate, at pH 4.5, 0.025 M. One barrel of the pipette contained DL-homocysteic acid at pH 8.5, 0.1 M, and another contained 2 M NaCl to control for current effects. In some experiments a solution of naloxone (Endo Laboratories), 0.027 M at pH 4.5, was used in one barrel.



When cells which were not firing spontaneously were fired by the application of homocysteic acid, enkephalin and morphine had pronounced depressant actions, suggesting that their effects were probably exerted on the cell membrane rather than presynaptically.

The effects of met- and leu-enkephalin could often be seen with low ejection currents (5 or 10 nA), and the degree of inhibition was similar for both compounds. Thus leu-enkephalin, compared with met-enkephalin, seems to be more effective in the brain than it was found to be at peripheral receptors¹. In most cases higher currents were required to obtain the same effects with morphine. A reliable comparison of potencies, however, cannot be made since the relative efflux rates of these substances are not known and only a method such as that described by Gent *et al.*² would enable them to be determined.

Effects on reticular neurones are of interest since they respond to noxious stimulation^{3,4}. These cells show a wide convergence of afferent input⁵, and in some of the present experiments reticular neurones were identified by their response to pressure or pinch over large areas of the body surface⁶. Met-enkephalin was tested on 22 of these neurones — (21 inhibited, 1 unaffected). Morphine inhibited all 21 and leu-enkephalin all 13 on which they were tested. Thus reticular neurones with somatic responses in the

medulla or caudal pons were consistently inhibited by all three substances.

The effects of morphine on a variety of systems can be antagonised by naloxone. This drug also blocks the inhibition of brain stem neurones by morphine⁶, and it has now been shown also to block the inhibitory actions of met-enkephalin on brainstem neurones in the rat⁷. In the cat naloxone has been reported to antagonise the analgesia produced by bulbar raphe stimulation⁸. It was therefore surprising to find that neither the effects of morphine nor those of met-enkephalin on brainstem neurones in the cat could be blocked by naloxone. Although naloxone was applied for periods up to 35 min with currents of up to 150 nA (means 10 min and 81 nA), no appreciable antagonism was observed in the 16 neurones on which it was tested.

This result might suggest that the action of morphine on the neurones (including reticular neurones) which were investigated in the present experiments is unrelated to its analgesic action. Unfortunately naloxone does not, however, seem to have been tested for effects on analgesia produced by morphine in the cat. (There is a widely held belief that morphine is not analgesic in the cat, but doses of 1–2 mg kg⁻¹ intravenously have been reported to produce analgesia⁹; this action occurs at a much lower dose when the injections are made into the IVth ventricle¹⁰, suggesting that it is produced by an action on the brainstem.)

To conclude, both met- and leu-enkephalin have inhibitory actions similar to those of morphine on neurones in the brainstem of the cat, but surprisingly none of these effects can be blocked by naloxone. Thus the receptors responsible for these inhibitory actions must be different from those in the rat brainstem, and those in peripheral systems which can be antagonised by naloxone.

We thank the MRC for support.

J. P. GENT

J. H. WOLSTENCROFT

Department of Physiology,
Medical School,
University of Birmingham,
Birmingham B15 2TJ, UK

Received March 17; accepted April 22, 1976.

- Hughes, J., *et al.*, *Nature*, **258**, 577–579 (1975).
- Gent, J. P., Morgan, R., and Wolstencroft, J. H., *Neuropharmacology*, **13**, 441–447 (1974).
- Casey, K. L., *Expl Neurol.*, **25**, 35–56 (1969); *Science*, **173**, 77–80 (1971).
- Wolstencroft, J. H., *J. Physiol., Lond.*, **174**, 91–108 (1964).
- Scheibel, M., Scheibel, A., Mollica, A., and Moruzzi, G., *J. Neurophysiol.*, **18**, 309–331 (1955).
- Bramwell, G. J., and Bradley, P. B., *Brain Res.*, **73**, 167–170 (1974).
- Bradley, P. B., Briggs, I., Gayton, R., and Lambert, L. A., *Nature*, **261**, 425–426 (1976).
- Oliveras, J. L., Redjemi, F., Guilbaud, G., and Besson, J. M., *Pain*, **1**, 139–145 (1975).
- McKenzie, J. S., and Beechey, N. R., *E.E.G. Clin. Neurophysiol.*, **14**, 501–519 (1962).
- Dey, P. K., and Feldberg, W., *J. Physiol. Lond.*, **248**, 40–41P (1975).

Reduction of possible hazards in the preparation of recombinant plasmid DNA

THE preparation of plasmid DNA requires the collection of plasmid-containing bacteria, usually by centrifugation of liquid cultures. This step is a potential source of laboratory contamination, especially when large volumes, requiring continuous flow centrifugation, are processed.

We report here that bacteria killed by treatment with chloroform¹ or with diethylpyrocarbonate² are as good a starting material for the purification of plasmid DNA as are untreated bacteria. For these experiments we have used a streptomycin-dependent strain of *Escherichia coli* as recipient in transfection experiments and as host for the amplification of plasmid DNA.

We tested several strains of streptomycin-dependent *E. coli* (N543, N530 and N523, obtained from H. G. Wittmann (see ref. 3)) for their susceptibility to transfection by PCRI DNA, a chimaeric plasmid obtained by inserting a fragment of PSC105 DNA, containing the kanamycin resistance marker, into colE1 DNA⁴ and deleting one of the two *EcoRI* restriction sites⁵.

E. coli N543 (a streptomycin-dependent derivative of *E. coli* AB 774, *rec*⁺ *F*⁻ *lac*⁻ *xyl*⁻ *gal*⁻ *thi*⁻ *arg*⁻ *his*⁻ *ile*⁻, derived from *E. coli* K) was found to be very suitable for transfection; about 5×10^5 transfectants per μg of PCRI DNA were obtained under transfection conditions essentially as described by Mandel and Higa⁶. This strain compares favourably with the commonly used recipient *E. coli* HB101 (obtained from H. Boyer) which gave values of 1.8×10^5 transfectants per μg (Table 1). Since the streptomycin-dependent strain is *rec*⁺ and contains the K restriction and modification system, it is as such not recommended for the cloning of chimaeric DNA. A clone of PCRI-containing *E. coli* N543 (*E. coli* N543 (PCRI)) was isolated and found to be *lac*⁻, auxotrophic, resistant to kanamycin to at least $1,000 \mu\text{g ml}^{-1}$ and absolutely dependent on streptomycin for colony formation on agar plates. Half-maximal growth rate in tryptone medium required

Table 2 Effect of diethylpyrocarbonate or chloroform on the viability of bacterial cultures

Treatment	Time of exposure	Viability after treatment		
Diethylpyrocarbonate	None	at pH 5.5 1.3×10^8	at pH 6.5 1.7×10^8	at pH 7.5 1.8×10^8
	15 s	0	0	ND
	30 s	0	0	<10
	45 s	0	0	0
	60 s	0	0	0
Chloroform	None	at pH 6.2 3.4×10^8		
	60 s	500		
	90 s	67		
	120 s	0		
	240 s	0		

E. coli N543 (PCRI) was prepared by transfecting *E. coli* N543 with PCRI DNA as described in the legend to Table 1. The bacteria were grown in 100 ml of tryptone medium to an A_{650} of 0.8–0.9 on the Beckman DB spectrophotometer. At this point the pH of the culture had fallen from 7.2 to 6.2. Samples were adjusted to the pH indicated and 0.3 vol% of diethylpyrocarbonate or 1 vol% of chloroform were added with vigorous stirring at 37 °C. At the times indicated 1-ml aliquots were removed and diluted with 9 ml of chilled tryptone medium adjusted to pH 8 with NH_4OH . Each sample was diluted where appropriate, and a 1-ml sample was immediately filtered through a Millipore HAWP 0.45 μm filter (25 mm diameter). The filters were placed face upward on McConkey agar containing $200 \mu\text{g ml}^{-1}$ streptomycin and $100 \mu\text{g ml}^{-1}$ kanamycin and incubated at 37 °C overnight.

about $150 \mu\text{g ml}^{-1}$ and maximal rate $500 \mu\text{g ml}^{-1}$ of streptomycin. The level of streptomycin-independent revertants in stocks prepared from clones was about 1 in 10^8 . Bacteria grown in streptomycin-containing tryptone broth continued to grow for some generations even after removal of the antibiotic, presumably until streptomycin-saturated ribosomes were diluted out of the bacteria.

A late log phase culture of *E. coli* N543 (PCRI) was vigorously stirred with 1 vol% of chloroform or 0.3 vol% of diethylpyrocarbonate at 37 °C. At appropriate intervals 1-ml samples were diluted into 9 ml of chilled tryptone medium, and the viable bacteria were immediately determined. No surviving bacteria could be detected in the equivalent of 1 ml of the original culture after 120 s of treatment with chloroform or after 15 s of treatment with diethylpyrocarbonate (Table 2). Diethylpyrocarbonate was equally effective at an initial pH of 5.5, 6.5 or 7.5.

For purification of plasmid DNA from untreated and killed *E. coli*, a log phase culture of chloramphenicol-treated ³²P-labelled *E. coli* N543 (PCRI) was divided into two. One part was collected immediately, whereas the other was vigorously shaken with diethylpyrocarbonate for 5 min at 37 °C. NH_4OH was added to pH 7.5 to accelerate the decomposition of excess reagent after which the cells were collected. Both bacterial pellets were then suspended, lysed with Triton X-100 and the plasmid DNA was purified by a modification of the method of Katz *et al.*⁷. The yields and the specific infectivities of PCRI DNA (form I) were essentially the same for both samples (Table 3). Treatment with diethylpyrocarbonate thus affected neither the extraction nor the biological activity of the plasmid DNA. Similar results were obtained by the purification method of Guerry *et al.*⁸ which we found more convenient for the preparation of large quantities of plasmid DNA. Treatment of *E. coli* N543 (PCRI) with diethylpyrocarbonate for 10 min or longer occasionally rendered the bacteria less susceptible to lysis with Triton X-100; however, lysis with 1% sodium dodecyl sulphate, as used in the method of Guerry *et al.*⁸ was always successful.

Bacteria killed with chloroform were also perfectly suitable starting material, both with regard to yield and specific infectivity of plasmid DNA (data not shown).

The simple killing step introduced before collection

Table 1 Transfection of various strains of *E. coli* by PCRI DNA

PCRI DNA from untreated bacteria	<i>E. coli</i> HB101 Colonies/aliquot	<i>E. coli</i> HB101 Colonies/ μg DNA	<i>E. coli</i> N543 Colonies/aliquot	<i>E. coli</i> N543 Colonies/ μg DNA
0 ng	0	—	0	—
5 ng	101	1.2×10^5	516	5.9×10^5
20 ng	630	1.8×10^5	1,720	4.9×10^5
PCRI DNA from bacteria killed with diethylpyrocarbonate				
0 ng	0	—	0	—
5 ng	83	0.95×10^5	405	4.7×10^5
20 ng	623	1.8×10^5	1,760	5.1×10^5

The procedure is based on that of Mandel and Higa⁶. *E. coli* HB101 and *E. coli* N543 were grown in tryptone medium to an A_{650} (Beckman DB) of about 1. In the case of the strain N543 the medium contained $200 \mu\text{g ml}^{-1}$ of streptomycin sulphate. The bacteria were collected by centrifugation for 15 min at 3,000 r.p.m. in a Sorvall HB4 rotor, resuspended in 0.5 vol of 0.05 M CaCl_2 and kept for 15 min at 0 °C. The bacteria were then centrifuged and resuspended in 0.3 vol (based on the original volume) of 0.05 M CaCl_2 . The viable cell count was $5 \times 10^8 \text{ ml}^{-1}$ for *E. coli* HB101 and 10^9 ml^{-1} for *E. coli* N543. To 50 μl of PCRI DNA solution (in 10 mM Tris-HCl (pH 7.5)–10 mM CaCl_2 –10 mM MgCl_2) 100 μl of the bacterial suspension were added. The suspension was incubated for 15 min at 0 °C, 2 min at 37 °C and 10 min at room temperature. 1 ml of tryptone medium (containing $200 \mu\text{g ml}^{-1}$ of streptomycin in the case of *E. coli* N543) was added and incubation was continued for 1 h at 37 °C. 0.2 ml of each sample was then spread on McConkey agar containing $100 \mu\text{g ml}^{-1}$ of kanamycin (and $200 \mu\text{g ml}^{-1}$ of streptomycin in the case of *E. coli* N543). Tryptone medium contained per litre: Bacto tryptone, 10 g; yeast extract (Difco) 1 g; glucose, 1 g; NaCl, 8 g; $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 0.294 g.

Table 3 Purification of PCRI DNA from native or diethylpyrocarbonate-killed *E. coli* N543 (PCRI)

Yield	a, From native <i>E. coli</i>		b, From diethylpyrocarbonate-killed <i>E. coli</i>	
	A_{260} units	Total c.p.m.	A_{260} units	Total radio activity (c.p.m.)
Step 1 (Crude DNA)	ND	1.4×10^9	ND	1.5×10^9
Step 2 (after CsCl density gradient)	0.38	2.8×10^6	0.42	3.0×10^6
Step 3 (after sucrose density gradient)	0.26	1.6×10^6	0.36	1.7×10^6

A 2-l flask containing 0.5 l of low-phosphate tryptone medium (final concentration of inorganic phosphate, 0.25 mM) with 200 $\mu\text{g ml}^{-1}$ of streptomycin was inoculated with *E. coli* N543 (PCRI) and 1 h later 10 mCi of carrier-free ^{32}P -phosphate were added. At an A_{650} (Beckman DB) of about 0.8, chloramphenicol was added to a final concentration of 0.17 mg ml^{-1} and incubation continued for 14 h at 37 °C. The culture was divided into two 250-ml portions. Portion *a* was chilled immediately. Diethylpyrocarbonate (0.75 ml) were added to portion *b*. After 5 min at 37 °C two drops of 1% phenol red and concentrated NH_4OH were added until the indicator turned red and the culture was chilled. A 1-ml aliquot of this suspension was diluted with 10 ml of medium, the bacteria were collected by centrifugation, resuspended in 0.2 ml tryptone medium and spread on a streptomycin-containing (200 $\mu\text{g ml}^{-1}$) McConkey agar. No colonies were observed after 24 h. The bacteria were collected by centrifugation, and plasmid DNA was extracted essentially by the method of Katz *et al.*⁷. The bacteria were suspended in 50 ml of 10 mM Tris-HCl (pH 8.0), 1 mM EDTA, centrifuged and resuspended in 10 ml of 25% sucrose, 50 mM Tris-HCl (pH 8.5). Lysozyme (12.5 mg) was added and after 5 min at 0 °C, EDTA (pH 8.5) was added to 50 mM. After 5 min at 0 °C 10 ml of a solution containing 0.1% Triton, 0.53 M EDTA, 0.05 M Tris-HCl (pH 8.5) were added. After 1 h at 0 °C the viscous lysates were cleared by centrifugation and extracted with one volume of phenol-cresol-hydroxyquinoline (500:70:0.5 v/v/w) (step 1 preparation). The nucleic acids were precipitated with ethanol, redissolved and centrifuged to equilibrium in a CsCl gradient containing ethidium bromide. The lower stained band was collected, extracted 3 times with 1 vol of isoamylalcohol, dialysed and precipitated with ethanol (step 2 preparation). The preparation was centrifuged through a 4.5-ml 5–23% sucrose gradient in 50 mM Tris-HCl (pH 7.5) for 90 min at 58,000 r.p.m. and 15 °C in the Spinco SW60 rotor. The single peak of ^{32}P -PCRI DNA (step 3 preparation) was located by Cerenkov counting. Analysis of the DNA by agarose gel electrophoresis showed the exclusive presence of form I PCRI DNA in both samples *a* and *b*.

eliminates one of the major sources of environmental contamination in work with potentially pathogenic *E. coli* and should greatly simplify the physical containment requirements which might otherwise be necessary for the processing of larger quantities of bacterial cells. The use of streptomycin-dependent strains of *E. coli* should further reduce the probability of unwanted propagation of plasmid-containing *E. coli*.

We thank Dr H. G. Wittmann for the streptomycin-dependent strains of *E. coli*, Dr H. Boyer for *E. coli* HB101, Dr Grunstein for the plasmid PCRI (ref. 5), Dr J. Carbon for unpublished information and Dr W. Arber for determining the restriction system in *E. coli* N543. This work was supported by a grant from the Schweizerische Nationalfonds.

C. WEISSMANN
W. BOLL

Institut für Molekularbiologie I,
Universität Zürich,
8049 Zürich, Switzerland

Received March 1; accepted April 7, 1976.

- Adams, M. H., *Bacteriophages* (Interscience, New York, 1959).
- Pauli, O., and Genth, H., *Z. Lebensm. Unters. Forsch.*, **132**, 216–227 (1966).
- Hasenbank, R., *et al.*, *Molec. gen. Genet.*, **127**, 1–18 (1973).
- Herschfeld, V., Boyer, H. W., Yanofsky, C., Lovett, M. A. and Helinski, D. R., *Proc. natn. Acad. Sci. U.S.A.*, **71**, 3455–3459 (1974).
- Covey, C., Richardson, D., and Carbon, J., *Molec. gen. Genet.* (in the press).
- Mandel, M., and Higa, A., *J. molec. Biol.*, **53**, 159–162 (1970).
- Katz, L., Kingsbury, D. T., and Helinski, D. R., *J. Bact.*, **114**, 577–591 (1973).
- Guerrey, P., LeBlanc, D. J., and Falkow, S., *J. Bact.*, **116**, 1064–1066 (1973).

Phosphorylation of yeast RNA polymerases

DNA-DEPENDENT RNA polymerase activities in eukaryotic cells change during developmental^{1,2} and physiological transitions^{3–5}. Some changes occur rapidly and in the absence of protein synthesis. They presumably involve modulation of the activity of pre-existing RNA polymerase molecules rather than alterations of the rate of synthesis or degradation of these molecules. This modulation of polymerase activity could be a consequence of covalent modification of enzyme subunits, the rapid turnover of a protein factor required for catalysis, alteration of the template properties of the chromatin or some combination of these. The first possibility is attractive. The activity of many enzymes is controlled in this manner⁶. Cycles of phosphorylation and dephosphorylation regulate the activity of the enzymes involved in glycogen metabolism (reviewed

in refs 9 and 10). Lee *et al.*¹¹ observed an increase in the specific activity of Rous sarcoma virus reverse transcriptase on incubation with protein kinase and ATP. Phosphorylation of β and β' subunits of *E. coli* RNA polymerase on infection by bacteriophage T7 has been implicated in the control of early transcription of T7 DNA¹². Several groups^{5,13,14} have proposed and presented preliminary evidence that phosphorylation of RNA polymerase is involved in regulation of polymerase activity in eukaryotes. Jungmann *et al.*⁵ and Martelo and Hirsch¹³ observed enhanced RNA polymerase activity when partially purified protein kinase, RNA polymerase and ATP were mixed. They observed the phosphorylation of proteins in their preparations and attributed the increase in polymerase activity to phosphorylation of the enzyme. They did not demonstrate phosphorylation of RNA polymerase polypeptides, and since the observed increase in RNA polymerase activity could also be a consequence of phosphorylation of contaminating acidic nuclear proteins which can stimulate polymerase activity^{15–17}, the cause of the enhanced polymerase activity is obscure. We report here the isolation of phosphorylated RNA polymerase I from yeast cells and the identification of the enzyme polypeptides phosphorylated *in vivo*. The same pattern of phosphorylation was obtained when highly purified RNA polymerase I was incubated with a yeast protein kinase preparation.

Yeast cells were grown in complete medium in the presence of ^{32}P -phosphate and collected in the late log phase of growth. RNA polymerase I was purified partially, mixed with unlabelled RNA polymerase I and subjected to sucrose gradient centrifugation. A significant fraction of the phosphorylated proteins sedimented with the enzymatic activity of the added RNA polymerase (Fig. 1a). The possible relationship of the phosphorylated polypeptides to the putative RNA polymerase I subunits was determined by resolution of the protein from the sucrose gradient fractions by polyacrylamide gel electrophoresis and subsequent autoradiography of the dried gel. Phosphorylated polypeptides migrated together with five of eleven polymerase I polypeptides (Fig. 1b and c). The 185,000, 44,000, 35,000 and 24,000-dalton polypeptides were relatively more labelled than the 20,000-dalton polypeptide. There was also a small amount of radioactivity which migrated with the dye-front and may have been RNA. These data indicate that RNA polymerase is phosphorylated *in vivo* and is a phosphoprotein as isolated by common procedures.

Purified yeast RNA polymerases can be phosphorylated

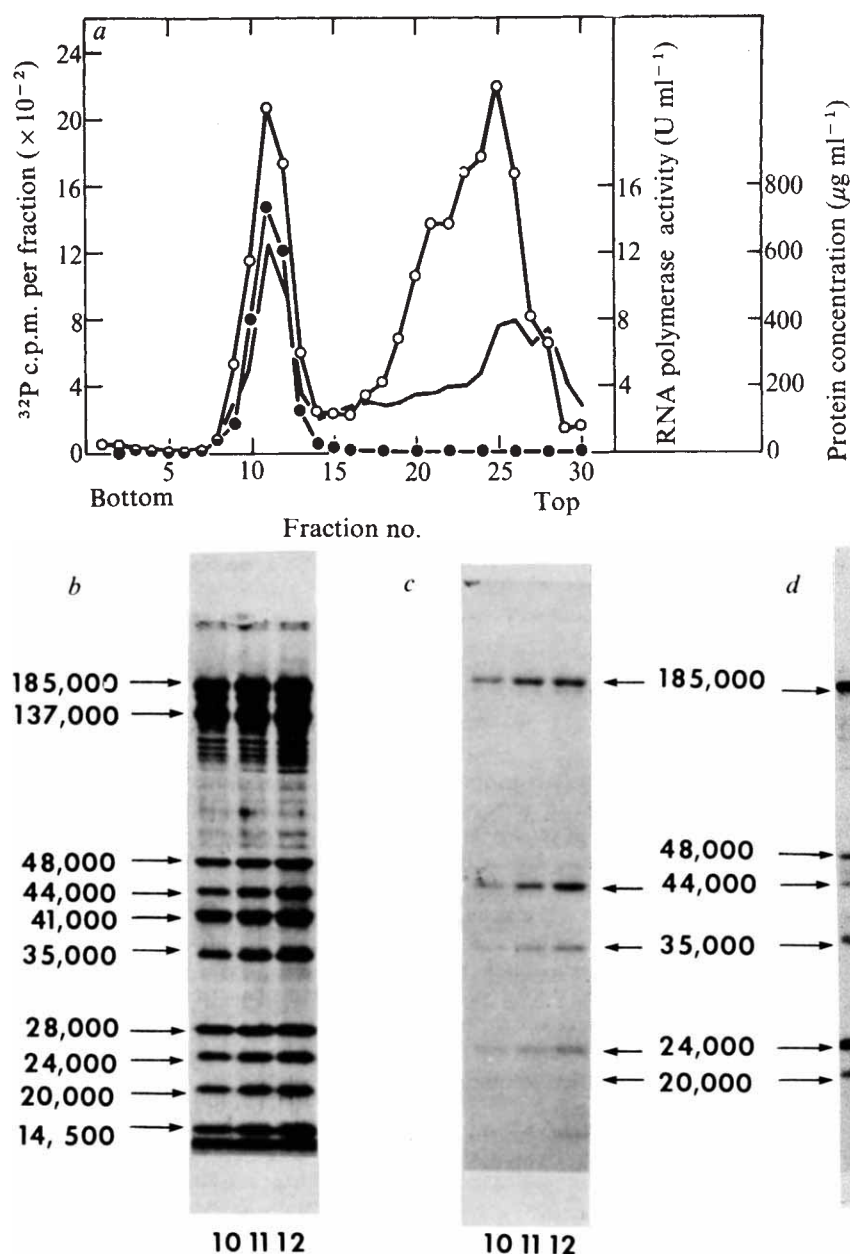


Fig. 1 Phosphorylation of yeast RNA polymerase I. ^{32}P -labelled RNA polymerase I was isolated from a prototrophic tetraploid strain of *Saccharomyces cerevisiae* (provided by Dr L. Hartwell) grown in complete medium (YEPD) 18 containing ^{32}P -phosphate ($5 \mu\text{Ci ml}^{-1}$). Cells (2 l) were collected in the late-log phase of growth (30 g , wet weight) and spheroplasts were prepared by the two-step procedure of Cabib 19 as described by Wintersberger *et al.* 20 . RNA polymerase I was extracted in the presence of 1 mM ATP and was purified to fraction 3 by a micro-adaptation of Valenzuela *et al.* 21 . The ^{32}P -labelled enzyme fraction was mixed with unlabelled polymerase I (fraction 4) 21 and protein was precipitated by dialysis against a saturated ammonium sulphate solution. The precipitate was collected by centrifugation ($10,000\text{g}$, 20 min) and dissolved in 15% glycerol, 0.2 M KCl, 0.1 mM EDTA, 6 mM 2-mercaptoethanol, 50 mM Tris-HCl, pH 7.9. The RNA polymerase I was further purified by centrifugation in a $5\text{--}20\%$ (w/v) sucrose gradient in 0.5 M KCl, 15% glycerol, 0.1 mM EDTA, 6 mM 2-mercaptoethanol, 50 mM Tris-HCl, pH 7.9 (SW56 rotor, $56,000 \text{ r.p.m.}$, 15 h , 4°C). Fractions were collected by tube puncture. *a*, Sucrose gradient profile of RNA polymerase I phosphorylated *in vivo*. RNA polymerase activity and protein concentration of each fraction were determined as before 21 . An aliquot of each gradient fraction was spotted on to a glass fibre filter (GF/C, Whatman). Filters were washed in 10% trichloroacetic acid (TCA). Contaminating ^{32}P -labelled nucleic acid was hydrolysed by boiling the filters in 5% TCA for 15 min . The filters were then washed twice more in 5% TCA, then in ethanol and dried. The ^{32}P -labelled protein was counted in Omnifluor (New England Nuclear)-toluene scintillant. *b*, Resolution of polypeptides by sodium dodecyl sulphate (SDS)-acrylamide gel electrophoresis of *in vivo* ^{32}P -labelled RNA polymerase. The polypeptides in each sucrose gradient fraction were resolved by 10% acrylamide slab-gel electrophoresis in the presence of 0.1% SDS 21 and stained with Coomassie blue. Buffers and solutions were prepared according to Laemmli 22 . The putative subunits of RNA polymerase I and assigned molecular weights are indicated by the arrows. *c*, Autoradiogram of RNA polymerase polypeptides labelled *in vivo* with ^{32}P . The stained gel prepared in (*b*) was dried and the ^{32}P -polypeptides were located

by autoradiography. *d*, *In vitro* phosphorylation of yeast RNA polymerase I subunits. The standard reaction mixture contained in 0.050 ml : $10 \mu\text{g}$ of purified RNA polymerase I (21), 0.2 U of yeast protein kinase (1 U of protein kinase activity defined as 1 nmol of ^{32}P transferred from ^{32}P -ATP to $200 \mu\text{g}$ of phosphatase in 10 min at 30°C), 60 mM Tris-HCl, pH 7.9, 5 mM MgCl_2 , 6 mM NaF, 90 mM KCl, 5 mM 2-mercaptoethanol, 0.010 mM ^{32}P -ATP (specific activity, $1,000 \text{ d.p.m. pmol}^{-1}$) and 7.5% glycerol. After incubation for 15 min at 30°C , the reaction was stopped by the addition of 0.005 ml of 5 mM ATP and subsequent transfer to 0°C . Phosphorylated polymerase was precipitated from 10% TCA and the labelled polypeptides were identified after SDS-acrylamide gel electrophoresis and autoradiography as described above. No phosphorylation was observed in the absence of either polymerase or protein kinase.

in vitro with a protein kinase which is purified together with RNA polymerase I but is resolved on sucrose gradient centrifugation. Whether this kinase is identical to those reported in yeast is unknown 23,24 . This enzyme uses either ATP or GTP as a phosphate donor, is not stimulated by cyclic AMP and will phosphorylate both acidic and basic phosphoprotein acceptors (G. I. B., unpublished results). The incubation of purified polymerase I and kinase resulted in the phosphorylation of the five polypeptides phosphorylated *in vivo* and, in addition, the $48,000$ -dalton polypeptide (Fig. 1*d*). The $185,000$ and $24,000$ -dalton polypeptides were phosphorylated to a greater degree than the other polypeptides. The protein kinase also phosphorylated other molecules in the polymerase preparation. These phosphorylated non-polymerase polypeptides did not migrate together with stained polypeptides in the gel and it is not known

whether they are a consequence of proteolysis of enzyme polypeptides or are contaminants of the polymerase preparation. Enzyme polypeptides of purified yeast RNA polymerases II and III were also phosphorylated by the protein kinase preparation (data not shown). The $33,500$ and $24,000$ -dalton polypeptides of polymerase II and the $53,000$, $24,000$ and $20,000$ -dalton polypeptides of polymerase III were phosphorylated. (The isolation and subunit composition of polymerases II and III have been reported elsewhere 25,26 .) Three of the phosphorylated polymerase I polypeptides (molecular weight $185,000$, $44,000$ and $35,000$) are unique to polymerase I. Two-dimensional polyacrylamide gel electrophoresis showed the $24,000$ -dalton polypeptide to be common to RNA polymerases I, II and III; the $20,000$ -dalton polypeptide seems common to polymerases I and III (P.V., unpublished). The phosphorylation of the $24,000$ -

dalton polypeptide of polymerases I, II and III and the 20,000-dalton polypeptide of polymerases I and III supports the contention that these are common subunits between the respective enzymes. Four of the phosphorylated polymerase I polypeptides (molecular weight 48,000, 44,000, 35,000 and 24,000) are readily dissociated from the enzyme and seem to be important for activity. Polymerase I has been isolated that lacks the 48,000 and 44,000-dalton polypeptides²⁶ or the 48,000 and 35,000-dalton polypeptides²⁷ and has reduced enzymatic activity on double-stranded DNA templates. Polymerase I lacking the 24,000-dalton polypeptide has also been isolated and is inactive on single- and double-stranded DNA and poly d(A-T) templates (P. V. and G. I. B., unpublished results).

In the phosphorylation experiments with purified polymerase I and protein kinase *in vitro*, only 0.5–1.0 mol of phosphate was bound to each mole of polymerase in optimal conditions. Since the data of the *in vivo* experiments indicate that there are at least five phosphorylation sites per enzyme, we presume that the purified polymerase is partially phosphorylated as isolated. The polypeptides of the purified polymerase molecules are phosphorylated to different extents in the *in vivo* and *in vitro* experiments. For example, the 44,000-dalton polypeptide is labelled more *in vivo* than *in vitro* while the reverse is found with the 24,000 and 20,000-dalton polypeptides. Thus there may be an equilibrium in the cell between phosphorylated and non-phosphorylated polymerase polypeptides. We observed that the 48,000-dalton peptide is phosphorylated only *in vitro*. One explanation for this result is that the enzyme molecule exists as a complex within the cell, such that the 48,000-dalton polypeptide is not accessible to protein kinase.

By modulating the charge density at a specific site in a polypeptide, phosphorylation and dephosphorylation could control the extent of interaction between enzyme polypeptides or between the enzyme and other molecules. In this way, phosphorylation of RNA polymerase could have a direct effect on transcription by altering the specific activity of the enzyme. It could also have an indirect effect on activity by modulating the association of polymerase with regulatory molecules or chromatin. Alternatively, phosphorylation may alter the turnover of the enzyme. We have begun to test the effects of phosphorylation on these aspects of polymerase function. In preliminary experiments, incubation of purified polymerase I with protein kinase and ATP or with bacterial alkaline phosphatase produced no effects on enzyme activity. These experiments are inconclusive so far because the isolated polymerase I is phosphorylated and the phosphatase may not remove the phosphate residues. We shall test the effects of phosphorylation on activity in standard assay conditions as well as the transcription of ribosomal genes in more physiological conditions, to elucidate the relationship between the phosphorylation of polymerase I and transcription.

This research was supported by a grant from the NIH. P.V. is a recipient of an international fellowship from the NIH. We thank G. Hager for purified yeast RNA polymerase II and III.

GRAEME I. BELL
PABLO VALENZUELA
WILLIAM J. RUTTER

Department of Biochemistry and Biophysics,
University of California, San Francisco,
California 94143

Received February 5; accepted March 30, 1976.

¹ Roeder, R. G., and Rutter, W. J., *Biochemistry*, **9**, 2543–2553 (1970).

² Roeder, R. G., *J. Biol. Chem.*, **249**, 249–256 (1974).

³ Gross, K. G., and Pogo, A. O., *J. Biol. Chem.*, **249**, 568–576 (1974).

⁴ Griswold, M. D., and Cohen, P. P., *J. Biol. Chem.*, **248**, 5854–5860 (1973).

⁵ Jungmann, R. A., Hiestand, P. C., and Schweppe, J. S., *J. Biol. Chem.*, **249**, 5444–5451 (1974).

⁶ Jensen, E. V., and DeSombre, E. R., *A. Rev. Biochem.*, **41**, 203–230 (1972).

⁷ Means, A. R., and O'Malley, B. W., in *Biochemistry of Cell Differentiation* (edit. by Paul, J.), 161–180 (University Park Press, Baltimore, 1974).

⁸ Rutter, W. J., Goldberg, M. I., and Perriard, J. C., in *Biochemistry of Cell Differentiation* (edit. by Paul, J.), 267–300 (University Park Press, Baltimore, 1974).

⁹ Rubin, C. S., and Rosen, O. M., *A. Rev. Biochem.*, **44**, 831–887 (1975).

¹⁰ Taborsky, G., in *Advances in Protein Chemistry*, **28** (edit. by Anfinsen, C. B., Edsall, J. T., and Richards, F. M.), 1–120 (Academic, New York, 1974).

¹¹ Lee, S. G., Miceli, M. V., Jungmann, R. A., and Hung, P. P., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 2945–2949 (1975).

¹² Zillig, W., et al., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 2506–2510 (1975).

¹³ Martelo, O. J., and Hirsch, J., *Biochem. biophys. Res. Commun.*, **58**, 1008–1015 (1974).

¹⁴ Rutter, W. J., Morris, P. W., Goldberg, M., Paule, M., and Morris, R. W., in *The Biochemistry of Gene Expression in Higher Organisms* (edit. by Polak, J. K., and Lee, J. W.), 89–104 (Australia and New Zealand Book Co., 1973).

¹⁵ Shea, M., and Kleinsmith, L. J., *Biochem. biophys. Res. Commun.*, **50**, 473–477 (1973).

¹⁶ Teng, C. S., Teng, C. T., and Allfrey, V. G., *J. Biol. Chem.*, **246**, 3597–3609 (1971).

¹⁷ Stein, G. S., Spelsberg, T. C., and Kleinsmith, L. J., *Science*, **183**, 817–824 (1974).

¹⁸ Link, G. R., *Meth. Enzymol.*, **17A**, 59–78 (1970).

¹⁹ Cabib, E., *Meth. Enzymol.*, **22**, 120–122 (1971).

²⁰ Wintersberger, U., Smith, P., and Letnansky, K., *Eur. J. Biochem.*, **33**, 123–130 (1973).

²¹ Valenzuela, P., Weinberg, F., Bell, G., and Rutter, W. J., *J. Biol. Chem.*, **251**, 1464–1470 (1976).

²² Laemmli, U. K., *Nature*, **227**, 680–685 (1970).

²³ Takai, Y., Yamamura, H., and Nishizuka, Y., *J. Biol. Chem.*, **249**, 530–535 (1974).

²⁴ Lerch, K., Muir, L. W., and Fischer, E. H., *Biochemistry*, **14**, 2015–2023 (1975).

²⁵ Valenzuela, P., Hager, G. L., Weinberg, F., and Rutter, W. J., *Proc. natn. Acad. Sci. U.S.A.*, **73**, 1024–1028 (1976).

²⁶ Hager, G. L., Holland, M. J., Valenzuela, P., Weinberg, F., and Rutter, W. J., in *RNA Polymerase* (edit. by Chamberlin, M., and Losick, R.), (Cold Spring Harbor Laboratory, New York, in the press).

²⁷ Huet, J., Buhler, J. M., Sentenac, A., and Fromageot, P., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 3034–3038 (1975).

Association of RNase H activity with yeast RNA polymerase A

EUKARYOTIC RNA polymerases were isolated as large multimeric protein complexes containing two high molecular weight subunits and a collection of smaller polypeptide chains^{1–3}. This structural complexity suggests a multifunctional system, organised around a core enzyme combined with specificity determinants and possibly other proteins involved in regulation of chromatin activity. The purpose of this report is to describe the association of a ribonuclease H activity with yeast RNA polymerase A. This nuclease activity, which specifically degrades RNA–DNA hybrids, is inseparable from the polymerase by various fractionation procedures.

In the experiment shown in Fig. 1a, highly purified RNA polymerase A from yeast cells was subjected to phosphocellulose chromatography. The fractions were tested for RNA polymerising activity as well as for their ability to degrade the homopolymer ³H–(rA)_n or the hybrid duplex ³H–(rA)_n·(dT)_n. As previously reported⁴, RNA polymerase A was separated into two enzymatic fractions, RNA polymerase A*, that lacked two polypeptide chains, and the complete enzyme A. A nuclease activity that degraded the (rA)_n moiety of the hybrid into acid-soluble material was also revealed. But no degradation of the ribohomopolymer (rA)_n occurred. A coincident distribution of polymerase and nuclease activities was found at the level of the two protein peaks (Fig. 1a). When RNA polymerase A from phosphocellulose was subjected to zone sedimentation through a 10 to 30% glycerol gradient, the nuclease cosedimented with the polymerase activity (Fig. 1b). The two peaks A and A*, from phosphocellulose, were further chromatographed separately on DEAE-cellulose. Again, RNA polymerase A (Fig. 1c), as well as A* (result not shown), were eluted in a sharp protein peak, simultaneously with the (rA)_n·(dT)_n degrading activity. Furthermore, the nuclease, when subjected to electrophoresis on polyacrylamide gel under non-denaturing conditions, also had the same electrophoretic mobility as RNA polymerase. Again the nuclease was recovered at the level of the protein bands corresponding to RNA polymerase A and A* (ref. 4) (result not shown). From the above observations, since the nuclease activity always copurified with RNA polymerase during ion exchange chromatography, zone sedimentation and electrophoresis, it was concluded that the nuclease was physically associated with the RNA polymerase molecule.

As already seen in Fig. 1a the presence of the hybridised $(dT)_n$ strand was required for degradation of the $(rA)_n$ polymer. Likewise, removal of the $(dT)_n$ moiety of the hybrid by DNase prevented the degradation of the ribopolymer strand (Table 1). The polymerase-associated nuclease activity therefore had the properties of a ribonuclease H (ref. 5). Degradation of hybridised RNA was not restricted to the synthetic homopolymer duplex $(rA)_n \cdot (dT)_n$. 3H -RNA-T7 DNA hybrid, synthesised with *Escherichia coli* RNA polymerase, was also used as substrate with similar efficiency. On the other hand, the enzyme was completely free of DNase activity, as tested with 3H -T7 DNA by sedimentation in alkaline sucrose gradient.

It should be stressed that the ratio of RNase H compared with polymerase activity, as seen in Fig. 1a, is misleading, because the degradation reaction was not linear with time. The reaction tended to plateau at about 40 to 50% completion. Kinetic measurements, in fact, indicated that, as

far as RNase H was concerned, the specific activity of A* was four times lower than that of the complete enzyme A. This suggested that the two polypeptides missing in A* are in some way involved in the degradation process. Experiments with yeast RNA polymerase B did not reveal the presence of similar nuclease activity in this enzyme. Therefore, it is unlikely that the three common subunits of A and B enzymes³ are implicated.

The requirements of the nuclease are summarised in Table 1. In the above experiments, the activity was assayed at pH 8 (Fig. 1a and b) or pH 6.5 (Fig. 1c). In fact, $(rA)_n \cdot (dT)_n$ degradation occurred within a broad range of pH, between pH 6 and pH 8. It required the presence of divalent cations, either Mg^{2+} (5 mM), or preferably Mn^{2+} (1 mM), for optimal activity, and was sensitive to the ionic strength of the reaction mixture (Table 1). In the presence of 100 mM KCl, the nuclease retained only 27% of its optimal activity. Among other compounds, inorganic pyrophosphate (0.1 mM), phosphate (0.1 mM) or α -amanitin, were neither stimulatory nor inhibitory.

Since the requirements for RNase H activity associated with the polymerase were optimally met by the conditions used for transcription, the question arose whether the nuclease activity was expressed during RNA synthesis. With $(rA)_n \cdot (dT)_n$ as template primer, ATP can be incorporated by yeast RA polymerase A, possibly through *de novo* chain initiation or by elongation of pre-existing RNA chains (result not shown). In a challenge experiment, $(rA)_n \cdot (dT)_n$ degradation was followed in the presence of added ATP or with a different nucleoside triphosphate. The results are summarised in Table 1. With 0.5 mM ATP, in conditions where RNA synthesis took place, RNase H activity was depressed to an insignificant level. By contrast,

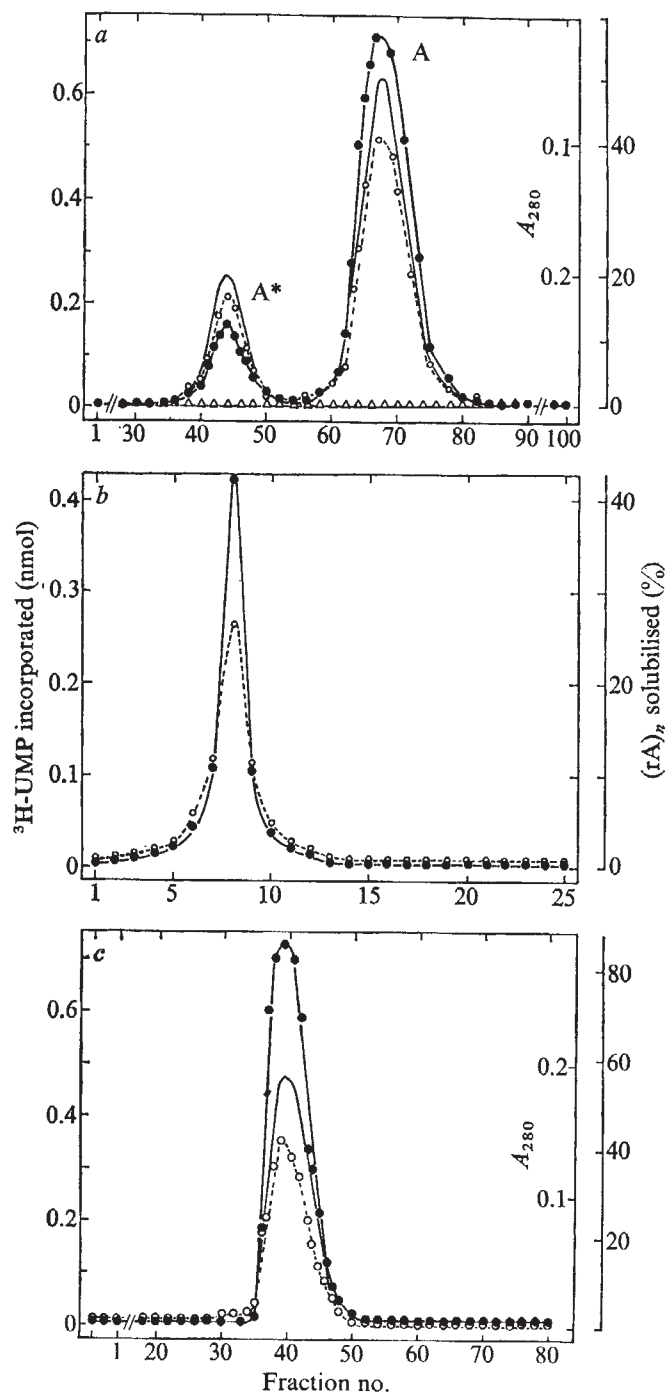


Fig. 1 Copurification of RNA polymerase and RNase H activities. *a*, Highly purified RNA polymerase A (5.8 mg of protein, glycerol gradient step) was chromatographed on phosphocellulose as previously described⁴, and the fractions (1 ml) were analysed for absorbance (—), RNA polymerase (●), RNase H (○) and $(rA)_n$ degrading (△) activities; *b*, RNA polymerase A from the phosphocellulose column (50 µg of protein, peak A) was loaded on top of 4.8 ml of a 10–30% glycerol gradient, in a buffer containing 0.02 M Tris-HCl pH 8, 0.5 mM EDTA, 0.01 M 2-mercaptoethanol, 0.05 M ammonium sulphate. The gradient was centrifuged for 18 h, at 40,000 r.p.m., in a SW65 rotor, at 2 °C. Fractions (0.2 ml) were collected from the bottom of the tube and assayed for RNA polymerase (●) and RNase H (○) activities. *c*, RNA polymerase A (3 mg of protein, peak A from phosphocellulose) was dialysed for 1.5 h at 4 °C against a buffer containing 0.02 M Tris-HCl, pH 8, 0.5 mM EDTA, 0.01 M 2-mercaptoethanol, 0.05 M ammonium sulphate and 30% glycerol (v/v), and applied to a DEAE-cellulose column (0.5 cm² × 12 cm) equilibrated with the same buffer containing 10% glycerol (v/v). The column was developed with 100 ml of a linear gradient from 0.05 M to 0.4 M ammonium sulphate in the above buffer. Fractions (1.4 ml) were collected and assayed for absorbance (—), RNA polymerase (●), and RNase H (○) activities. RNA polymerase was assayed in 0.1 ml mixtures containing 0.06 M Tris-HCl, pH 8, 1 mM dithiothreitol, 5 mM $MgCl_2$, 1 mM each of ATP and 3H -UPT (25 c.p.m. pmol⁻¹), 3 µg of d(A-T)_n and 10 µl from the various fractions. After 30 min incubation at 30 °C, radioactivity of acid-precipitable material was measured. RNase H activity was assayed in 0.1 ml mixtures containing 3H -(rA)_n·(dT)_n as substrate (0.22 nmol of (rA)_n; 36 c.p.m. pmol⁻¹ of polymerised nucleotide) 0.06 M Tris-HCl, pH 8.1 dithiothreitol, 5 mM $MgCl_2$ and 10 µl sample from the various fractions. In the case of the DEAE cellulose column, the assay was conducted at pH 6.5 with 0.025 M Tris-maleate buffer and with 0.05 mM $MgCl_2$. After 1 h incubation at 37 °C, 100 µg serum albumin (50 µl) were added as carrier, followed by 15 µl of 50% trichloroacetic acid; the precipitate was discarded by low speed centrifugation, and radioactivity of acid-soluble material was measured. The results are expressed as percentage of radioactivity released from the hybrid, relative to the input. (rA)_n-degrading activity was assayed in the above conditions, using the same amount of non-hybridised 3H -(rA)_n as substrate.

CTP or GTP, which were non-complementary nucleotides in this system, did not significantly affect degradation of the (rA)_n strand. Only UTP, which was complementary to the ribopolymer strand, was somewhat inhibitory, although to a lesser extent than ATP. To demonstrate in addition that inhibition caused by ATP was linked to the occurrence of the polymerisation reaction, the effect of α - β methylene ATP (0.5 mM) was investigated. This analogue of ATP cannot be used for chain elongation, but can support pyrophosphate exchange⁶, as initiator nucleotide, in the presence of d(A-T)_n and UTP (ref. 7). As expected, this analogue did not influence the RNase H activity. The absence of degradation during RNA synthesis probably reflected the processivity of the polymerisation reaction. This also indicated that the nuclease remained associated to the polymerase molecule during transcription.

Table 1 Requirements for (rA)_n-(dT)_n degradation

Addition or omission	Acid-solubilised radioactivity (c.p.m.)	% Hybrid degradation
Complete system	4,100	25.5
-MgCl ₂ +MnCl ₂ (1 mM)	7,800	49
-MgCl ₂	200	1
+10 μ g DNase at zero time	213	1
+20 μ g α -amanitin	3,880	24.5
+KCl (50 mM)	1,942	12
+KCl (100 mM)	1,130	7
+PP _i (0.1 mM)	4,120	25.5
+P _i (0.1 mM)	4,100	25.5
+ATP (0.5 mM)	197	1
+UTP (0.5 mM)	2,480	15.5
+GTP (0.5 mM)	3,340	21
+CTP (0.5 mM)	3,600	22.5
+ α - β methylene ATP (0.5 mM)	3,500	22

RNA polymerase A (1 μ g, from the DEAE-cellulose column shown in Fig. 1) was incubated, for 1 h at 37 °C, in 0.2 ml of a mixture containing ³H-(rA)_n-(dT)_n (448 pmol of polymerised ribonucleotides; 36 c.p.m. pmol⁻¹), 0.06 M Tris-HCl, pH 8.1, 5 mM MgCl₂ and 1 mM dithiothreitol. Radioactivity of acid-soluble material was measured as described in Fig. 1.

Nuclease activities were found physically associated with bacterial and phage-coded DNA-dependent RNA polymerases⁸. Mölling *et al.*⁹, also discovered the RNase H activity of viral reverse transcriptase. This, however, is the first time that a nuclease activity (clearly distinct from the reversal of polymerisation through pyrophosphorolysis) is found associated with a DNA-dependent RNA polymerase. The existence of this RNA polymerase-associated RNase H has significant implications in the consideration of the rôle of DNA-RNA hybrids in transcription. It is possible that RNA-DNA hybrids are involved in chromatin structure, perhaps regulating its template activity. Since yeast RNA polymerases seem to prefer unpaired DNA structures to initiate^{10,11}, transient RNA-DNA hybrids could also be formed at the initiation sites¹¹ and removed by the RNase H. Among other possibilities one could also consider a proof reading function for the polymerase. Ascribing a definitive physiological rôle to this nuclease activity will finally depend on genetic evidence.

JANINE HUET
FRANCOISE WYERS
JEAN-MARIE BUHLER
ANDRE SENTENAC
PIERRE FROMAGEOT

Service de Biochimie,
Département de Biologie,
CEN Saclay, BP No. 2,
91190—Gif-sur-Yvette,
France

- Chambon, P., *A. Rev. Biochem.*, **44**, 613-638 (1975).
- Sklar, V. E. F., Schwartz, L. B., and Roeder, R. G., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 348-352 (1975).
- Buhler, J. M., Iborra, F., Sentenac, A., and Fromageot, P., *J. biol. Chem.*, **251**, 1712-1717 (1976).
- Huet, J., Buhler, J. M., Sentenac, A., and Fromageot, P., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 3034-3038 (1975).
- Hausen, P., and Stein, H., *Eur. J. Biochem.*, **14**, 278-283 (1970).
- Krakow, J. S., and Fronk, E., *J. biol. Chem.*, **244**, 5988-5993 (1969).
- Sentenac, A., Dezelee, S., Iborra, F., Buhler, J.-M., Huet, J., Wyers, F., Ruet, A., and Fromageot, P., in *RNA Polymerases* (edit. by Chamberlin, M., and Losick, R.), (Cold Spring Harbor Monograph series, in the press).
- Kornberg, A., *Science*, **163**, 1410-1417 (1969).
- Mölling, K., Bolognesi, D. P., Bauer, H., Busen, W., Plassmann, H. W., and Hausen, P., *Nature new Biol.*, **234**, 240-243 (1971).
- Frederick, E. W., Maitra, U., and Hurwitz, J., *J. biol. Chem.*, **244**, 413-424 (1969).
- Dezelee, S., Sentenac, A., and Fromageot, P., *J. biol. Chem.*, **249**, 5971-5977 (1974).

Control of 5S RNA synthesis in *Xenopus laevis*

PREVIOUS reports¹⁻³ have shown that somatic cells of *Xenopus laevis* synthesise a 5S RNA which differs in sequence from 5S RNA synthesised by oocytes. In none of these reports is there any indication of ovary-type sequences being synthesised by somatic cells. Moreover, although the oocyte 5S sequences are heterogeneous, the somatic sequence is essentially homogeneous. We report here that *X. laevis* liver cells labelled in organ culture do not detectably synthesise ovary-type 5S RNA sequences but find that a transformed tissue culture cell line, derived initially from somatic kidney cells, synthesises significant (10-20%) amounts of ovary-type 5S RNA sequences. Taken together with the observation that some 5S genes are translocated to a novel site adjacent to the nucleolar organiser in these cells⁴ this result is important for an understanding of the nature of differential control of 5S gene transcription in *X. laevis*.

Total 5S RNA labelled with ³²PO₄ was prepared from either *Xenopus* liver cells or a permanent tissue culture cell line as described in the legend to Fig. 1. Oligonucleotide maps from complete RNase T₁ and RNase A digests of 5S RNA were made as described previously⁵ and the molar yields estimated. In liver, yields of radioactivity in regions of the chromatogram corresponding to the expected positions of oocyte-specific oligonucleotides were also counted to set an upper limit on the amount of oocyte type 5S sequence transcribed. In all cases blank areas of chromatogram were counted to provide background radioactivity estimates which were subtracted from the count rates of oligonucleotides. The results (Table 1) show that the three major oocyte-specific T₁ products (X₁, X₂ and X₃) and the one specific RNase A product (X₁) are all present with yields of 10-20% in 5S RNA from tissue culture cells but are not detectable above background in 5S RNA

Table 1 Molar yields of oligonucleotides from RNase T₁ and RNase A digests of 5S RNA from liver cells, tissue culture cells and oocytes of *X. laevis*

Nucleotide spot number*	Liver	Cell type		Oocyte*
		TC1†	TC2†	
RNase T ₁ products				
14 CCCG	1.03	0.83	0.92	0.09
17 ACCG	1.09	0.96	0.97	0.50
24 CCUG	1.83	2.47	2.23	2.90
51a CCAAG	1.07	0.77	0.90	0.27
53 AUCUCG	1.91	1.85	1.89	0.98
52 UACUUG	0.93	0.78	0.81	0.07
X ₁ AUCUCAG	0.01	0.23	0.10	0.70
X ₂ UACCUUG	0.01	0.19	0.13	1.06
X ₃ AUACAG	0.01	0.18	0.08	0.72
RNase A products				
107 GGAAGC	1.12	0.76	0.91	0.18
X ₁ AGAAGC	0.01	0.21	0.08	0.87

* Oocyte yields and spot numbers are those given in ref. 3.

† TC1 and TC2 refer respectively to tissue culture cells maintained by continuous passaging for 3 yr or in liquid nitrogen for 5 yr before analysis. Yields are the mean values for two determinations on two independent RNA preparations.

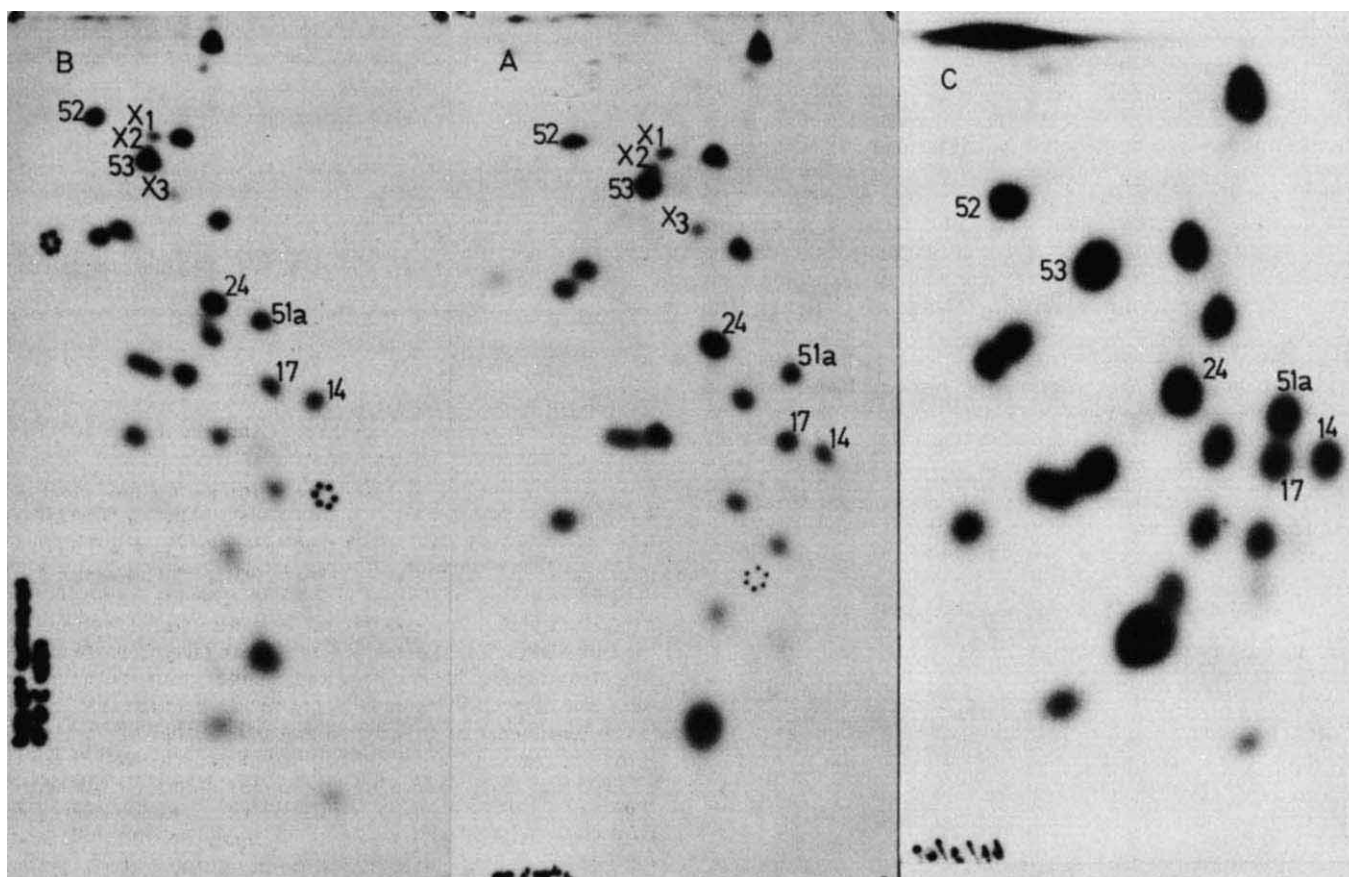


Fig. 1 Two-dimensional separation of RNase T_1 digestion products from ^{32}P 5S RNA labelled by tissue culture cell line 1 (A), tissue culture cell line 2 (B), and liver cells maintained in organ culture (C). Both tissue culture cell lines were derived from that established by Dr K. A. Rafferty in 1968 from adult *X. laevis* kidney. Tissue culture cell line 1 had been continuously grown over 3 yr before analysis, whereas tissue culture cell line 2 was derived from an ampoule of cells stored in liquid nitrogen for 5 yr before analysis. Tissue culture cells were grown to one-third confluence in Eagle's MEM plus 10% foetal bovine serum and diluted by 20% with water. They were labelled with carrier-free

^{32}P -orthophosphate (0.2 mCi ml^{-1}) for 3 d in the same medium containing one-third normal phosphate. Livers were removed from immature animals and rinsed well to remove as much blood as possible. They were pressed through a sterile 1-mm steel mesh. The small lumps of tissue ($\approx 1\text{-mm}$ cubes) were washed 5 times with fresh culture medium to remove broken cells and cultured for 48 h with carrier-free ^{32}P -orthophosphate (0.2 mCi ml^{-1}) in Eagle's MEM modified as above and containing one-third normal phosphate. 5S RNA was extracted, purified, digested with T_1 RNase and chromatographed as described previously³.

from liver cells. Further sequence analysis on these cell line products proved that they are identical with the sequences previously reported for oocyte oligonucleotides migrating to the same position on the chromatograms³.

We interpret these results as indicating a breakdown in the tissue culture line of the rigorous repression of oocyte-type 5S RNA transcription normally seen in somatic cells. The *Xenopus* kidney cell line is a transformed heteroploid cell line which has been maintained in culture for many years and has undergone considerable karyotype evolution. Nevertheless, a sample of the same cell line which had been kept in liquid nitrogen for several (five) years before analysis also produced oocyte-type 5S RNA sequences indicating that the event giving rise to leakiness of oocyte 5S gene repression must have been an early event in the establishment of this line, and is a constant and heritable change in the expression of 5S genetic material.

Genes coding for 5S RNA are present in many copies (24,000) in the genome⁶ and are localised in clusters at the ends of the long arms of most of the chromosomes in *X. laevis*⁴. In the normal karyotype, no 5S genes are detected in the region of the nucleolar organiser; however, *in situ* hybridisation of 5S cRNA to chromosomes of the same tissue culture cell line as used in these sequencing studies reveals significant 5S genetic material in this position. Morphologically, also, there seems to have been addition of chromosome material distal to the nucleolar organiser, presumably by translocation, in these

cells. Since the *in situ* technique will not distinguish between oocyte and somatic genes, it is not certain that this translocation involves oocyte-type 5S genes. Although normal liver cells must differ in many other ways from cells of the cultured cell line it is not easy to relate any of these differences in growth rate, differentiated epigenetic or genetic properties specifically to the control of 5S RNA synthesis. The simplest explanation for the both the qualitative and quantitative (Table 2) correlation between translocation of a single gene cluster and loss of complete repression is that the two events are causally related. If this is the case we can conclude that oocyte-type 5S genes are normally repressed in somatic cells not because there is any lack of polymerases or factors required for their transcription or because of the presence of specific diffusible inhibitors of transcription but because they are located in regions of the genome, such as heterochromatic regions, not normally available for transcription in somatic cells. By translocation to a region, such as the nucleolus organiser, which is normally expressed in somatic cells these oocyte 5S genes are now available for transcription. Assuming this interpretation is correct, we conclude that control of oocyte 5S gene repression in somatic cells is not an inherent property of the 5S DNA itself but of the chromosome region in which the 5S DNA is located. It cannot be due to diffusible factors (inducers or repressors) which interact with the 5S genes themselves since, in the absence of a higher level of control, such factors would operate at all points in a nucleus, and position would not affect their expression.

Table 2 Average relative transcription rates per gene for 5S RNA in three different *X. laevis* cell types

Total number of 5S genes	24,000	Total number of gene clusters	18
Average number of genes per cluster	1,340		
Number of somatic genes*	1,340		
Number of oocyte-type genes*	22,660		
Cell type	Average relative transcription rate per gene†		
	Somatic gene	Oocyte-type gene	
Oocyte	1 (10%)	0.53 (90%)	
Liver cell	1 (99%)	0.0007 (1%)	
Tissue culture cell a‡	1 (85%)	0.01 (15%)	
b	1 (85%)	0.35 (15%)	

* Number of somatic genes per haploid genome calculated assuming somatic genes are at a single cluster of average size, the rest of the genes are assumed to represent oocyte genes.

† The average relative transcription rate per gene is calculated by dividing the proportion of each type of 5S RNA by the appropriate number of genes and putting the somatic rate equal to 1. We assume no difference in stability of the two types of 5S RNA.

‡ Average relative rate of transcription in tissue culture cells assumes either (a) all oocyte genes are potentially active or (b) only one oocyte gene cluster (of average size) is active.

We thank Drs Ken Jones and R. Williamson for providing tissue culture cells, the SRC for support and Dr Ed Southern for critically reading the manuscript.

PETER J. FORD
TOM MATHIESON

Department of Molecular Biology,
University of Edinburgh, King's Buildings,
Mayfield Road, Edinburgh EH9 3JR, UK

Received January 19; accepted April 29, 1976.

¹ Brownlee, G. G., Cartwright, E., McShane, T., and Williamson, R., *FEBS Lett.*, **25**, 8–12 (1972).

² Wegnez, M., Monier, R., and Denis, H., *FEBS Lett.*, **25**, 13–20 (1972).

³ Ford, P. J., and Southern, E. M., *Nature new Biol.*, **241**, 7–10 (1973).

⁴ Pardue, M. L., Brown, D. D., and Birnstiel, M. L., *Chromosoma*, **42**, 191–203 (1973).

⁵ Brown, D. D., and Weber, C. S., *J. molec. Biol.*, **34**, 661–680 (1968).

Proton magnetic resonance study of the hydration of glucose

THE solvation of sugars by water has not been extensively studied in spite of its biochemical importance. Relaxation rates for ¹⁷O in monosaccharides have been used to study solvation^{1,2}, the revised² estimate of the solvation number being 5±1 water molecules for D-glucose and 2.5±1 for D-ribose at 5 °C, in agreement with 6±1 and 2.5±1 (respectively) derived from dielectric relaxation amplitudes. Again, compressibility data for aqueous sugars have been used³ to derive solvation numbers for monosaccharides, with results ranging from 2.5 to 4.0 at 25 °C. One limiting factor in nuclear magnetic resonance (NMR) studies of aqueous sugars is that the rate of proton exchange is fast and no separate resonances from the hydroxyl protons appear. We now report that at low temperatures the exchange rate becomes sufficiently slow in the pH range 4 to 8 to enable us to detect resolved resonances from the individual hydroxyl protons. The results are interpreted in terms of strong hydrogen bonding between water and glucose, the effective solvation number being at least 10.

At room temperature the NMR spectra of aqueous sugars have only a single peak for water and sugar-OH protons, because of rapid proton exchange. On cooling to about 0 °C, however, separate peaks for all the sugar-OH protons are resolved (Fig. 1). Typical data for D-glucose are included in Table 1. (See Fig. 2a for the structure and labelling

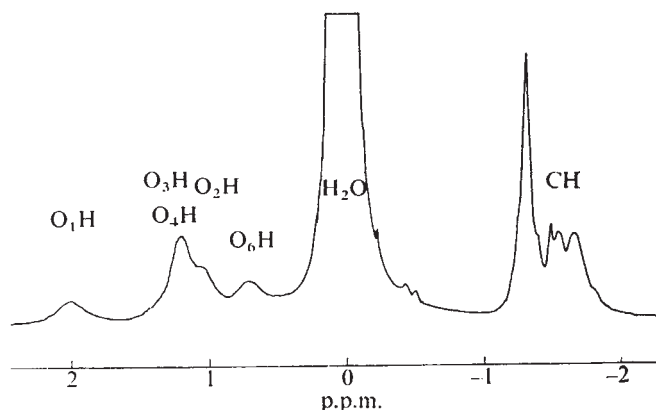


Fig. 1 100-MHz proton magnetic resonance spectrum of 2-molal α-D-glucose in water at -6 °C.

for β-D-glucose.) The following significant facts emerge: proton exchange with water is relatively slow for all hydroxyl protons. All ROH protons resonate down-field of the water protons, the anomeric hydroxyl proton resonance being considerably down-field of the remainder. All resonances shift up-field on heating, the shifts being comparable with that for water protons. They also broaden, the anomeric hydroxyl protons being most temperature sensitive, especially that for α-D-glucose. The effect of added basic cosolvents is, in general, similar to that on the water protons, comprising a relatively rapid up-field shift (Fig. 3).

Consider the solvated unit (Fig. 2b), in which two water molecules (also hydrogen bonded to other water molecules) are bonded to an OH group of a sugar molecule. Absence of (H₂O)α would give a (O-H)_{free} group which would resonate at very high field values, so (H₂O)α must be present. Furthermore, (H₂O)β has a marked strengthening effect on the hydrogen bond to (H₂O)α, and its absence would shift the O-H resonance to a value up-field of the pure water value.

The results can be compared with those for aqueous alcohols. In these systems, separate hydroxyl proton resonances are not usually resolved in the water-rich region, but extrapolation of the shifts observed in the alcohol-rich regions leads to the conclusion that hydrogen bonding

Fig. 2 a, β-D-glucose; b, solvated unit; c, concerted proton migration; d, example of other structures likely to form as a result of continued proton transfer.

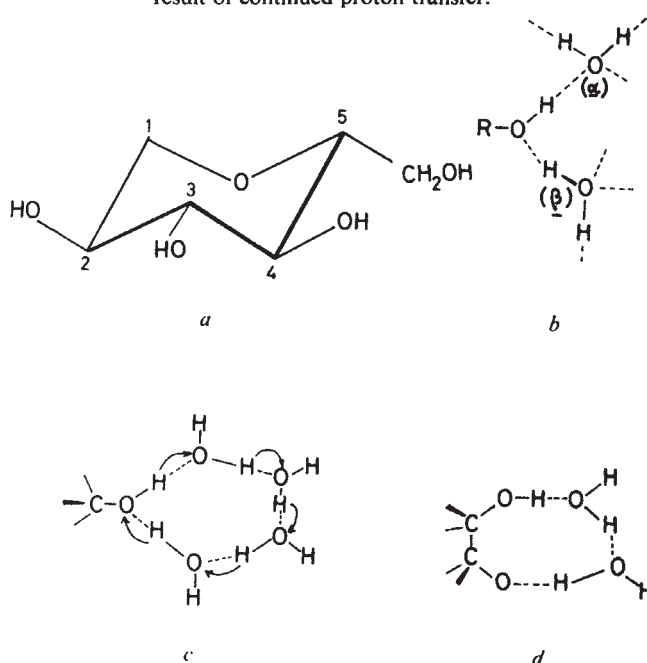


Table 1 Hydroxyl proton chemical shifts and temperature coefficients at 273 K

α -D-glucose (1 molal)			β -D-glucose (1 molal)		
Proton	Shift* (p.p.m.)	Temperature coefficient (p.p.m. °C ⁻¹)	Proton	Shift* (p.p.m.)	Temperature coefficient (p.p.m. °C ⁻¹)
O ₆ H	-0.73	0.016	O ₆ H	-0.96	0.011
O ₂ H	-1.05	0.021	O _{2, 3, 4} H	-1.27	0.018
O _{3,4} H	-1.23	0.016		-1.43	0.018
O ₁ H	-2.05	0.011	O ₁ H	-2.87	0.013
H ₂ O§	0	0.010	H ₂ O	0	0.011

*Shift is down field from the water resonance at 0 °C.

†Temperature coefficient of MeOH is 0.009 p.p.m. °C⁻¹.

‡Assignment based on that for DMSO solutions⁵ and the shifts obtained from aqueous DMSO (see Fig. 2a for numbering).

§Temperature coefficient of pure water is 0.0103 p.p.m. °C⁻¹.

between water and, for example, methanol, is actually more efficient than that between methanol molecules in the pure liquid⁴. The extrapolated resonance values are similar to those now reported for sugar hydroxyl protons. Thus, since it is generally agreed that methanol molecules in the pure liquid are nearly all doubly hydrogen bonded, this must also be true of methanol molecules in water and thus also of glucose in water.

The large down-field values for the anomeric hydroxyl

protons are in accord with their relatively high acidities, as are their greater linewidths on heating. One of the most striking features of our results is the fact that proton exchange with water is relatively slow compared with simple alcohols and also with glycol and glycerol. We have found that other cyclic monosaccharides also give resolved hydroxyl proton resonances below 0 °C in aqueous solution, and thus suggest that the controlling factor is the cyclic conformation. This is supported by the fact that the exocyclic hydroxyl proton exchanges more rapidly than any of the others. We tentatively propose the following argument to explain this observation.

Proton transfers of this type are acid-base catalysed, but in neutral solutions another, less efficient, process is probably involved, in which there is no charge separation. This can only come about by a form of concerted proton migration as depicted in Fig. 2c. Such structures are, of course, perfectly reasonable as fortuitous events in aqueous solution but, once formed, proton switching (as indicated by the arrows) will be facile. The lower the probability of such structures, the lower the rate of uncatalysed proton transfer. For (CH₃)₃COH, which does give a resolved O-H resonance at low temperature in water (our unpublished results), such structures are inhibited for steric reasons. This will also be true to some extent for the sugar hydroxyl groups, but another significant factor may be the tendency to form structures such as that in Fig. 2d, which cannot facilitate exchange, and which will compete with c and thus reduce the probability of d being formed. We suggest that these factors taken together explain the relatively slow proton exchange observed for sugar molecules.

The effect on the hydroxyl proton resonance of increasing the sugar concentration closely resembles that of adding methanol to the system, and the shifts are small. Dimethylsulphoxide and methylcyanide give rise to marked high-field shifts, because they remove the two hydrogen-bonded water molecules in turn, and only take the place of the one bonded to the hydroxyl proton (Fig. 2). Also, especially for methylcyanide, the resulting hydrogen bonds (ROH—NCCH₃) are relatively weak.

These results establish that, as anticipated, each hydroxyl group generally forms at least two hydrogen bonds to water.

We thank the MCR for a fellowship to J.M.H.

JOHN M. HARVEY
MARTYN C. R. SYMONS

Department of Chemistry,
The University,
Leicester LE1 7RH, UK

RICHARD J. NAFTALIN
Department of Physiology,
University of London King's College,
Strand, London WC2R 2LS, UK

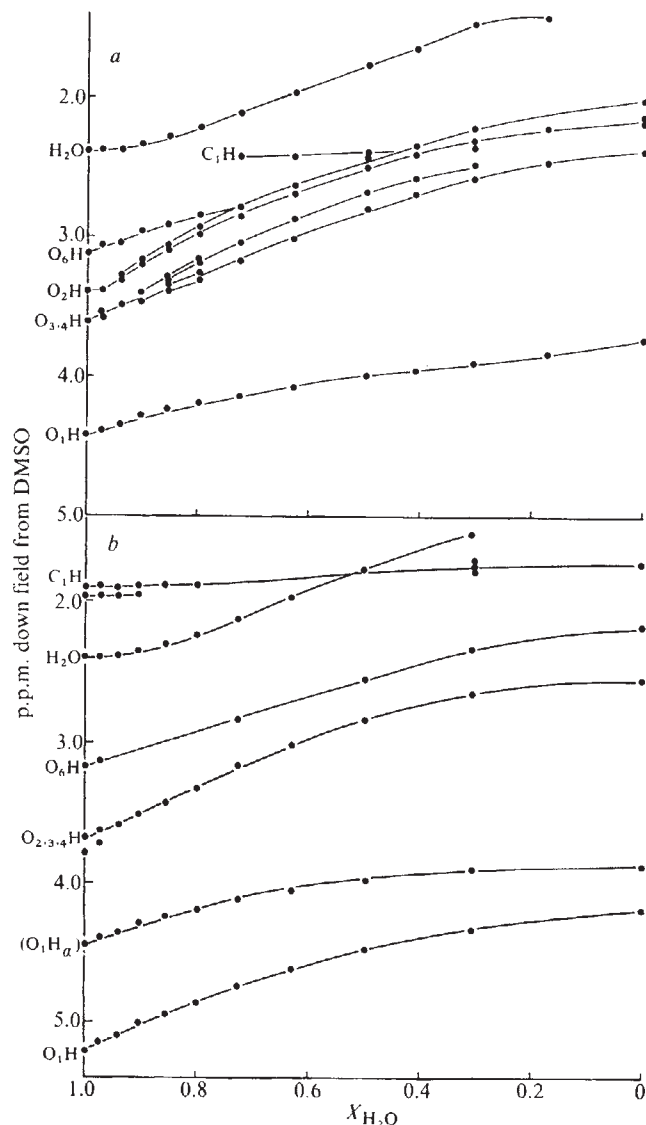


Fig. 3 Proton magnetic resonance of hydroxyl protons of α -D-glucose (a) and β -D-glucose (b) in water-dimethylsulphoxide mixtures at -6 °C. Hydroxyl proton shift against mole fraction of water. All shifts measured down field from methyl proton resonance of dimethylsulphoxide. Values were obtained using 1-molal glucose solutions (2-molal for pure dimethylsulphoxide).

Received January 22; accepted April 2, 1976.

- Tait, M. J., Suggett, A., Franks, F., Ablett, S., and Quickenden, P. A., *J. Solut. Chem.*, **1**, 131-151 (1972).
- Franks, F., Reid, D. S., and Suggett, A., *J. Solut. Chem.*, **2**, 99-113 (1973).
- Shio, H., *J. Am. chem. Soc.*, **80**, 70-73 (1958).
- Kingston, B., and Symons, M. C. R., *J.C.S. Faraday II*, **69**, 978-992 (1973).
- Perlin, A. S., *Can. J. Chem.*, **44**, 539-550 (1966).

matters arising

Experimental analysis of seismonastic movement in the sensitive plant

I HAVE studied the irritable reaction to stimuli of the pulvinated leaves of the "sensitive plant", *Mimosa pudica* L. using filming techniques similar to those described by Charnley *et al.*¹. Plants were grown in conditions very similar to theirs. Lighting was controlled strictly to a sequence of 14 h light ($5,600 \text{ erg cm}^{-2} \text{ s}^{-1}$) followed by 10 h darkness and the temperature was kept in the range $26.5 \pm 0.5^\circ \text{C}$. I report here an analysis of the factors influencing the recovery of the primary pulvinus.

The behaviour of the leaf after a shock can be divided into two phases. First, the collapse lasts a few seconds, during which the pinnules fold together in pairs and the primary pulvinus sinks. Second, the recovery lasts about 1 h during which many workers have observed oscillatory behaviour¹⁻³. A typical case is shown in Fig. 1.

For recovery, the curves of the variations of α against time can differ so much from leaf to leaf that many factors must influence turgor in the parenchymatous cells of the pulvinus. At least three factors are known to induce change in turgor in these cells—diurnal periodicity, phototropism and geotropism.

The first factor manifests itself in variations with time of day of seismonastic sensitivity⁶ and unstimulated angular position^{3,7-9}. These variations are small during the time around the middle of the period of illumination.

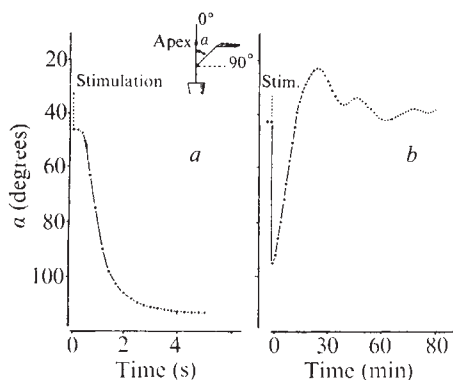


Fig. 1 Example of the time course in both phases of the seismonastic movement displayed by the primary pulvinus of the sensitive plant. *a*, Collapse; *b*, recovery.

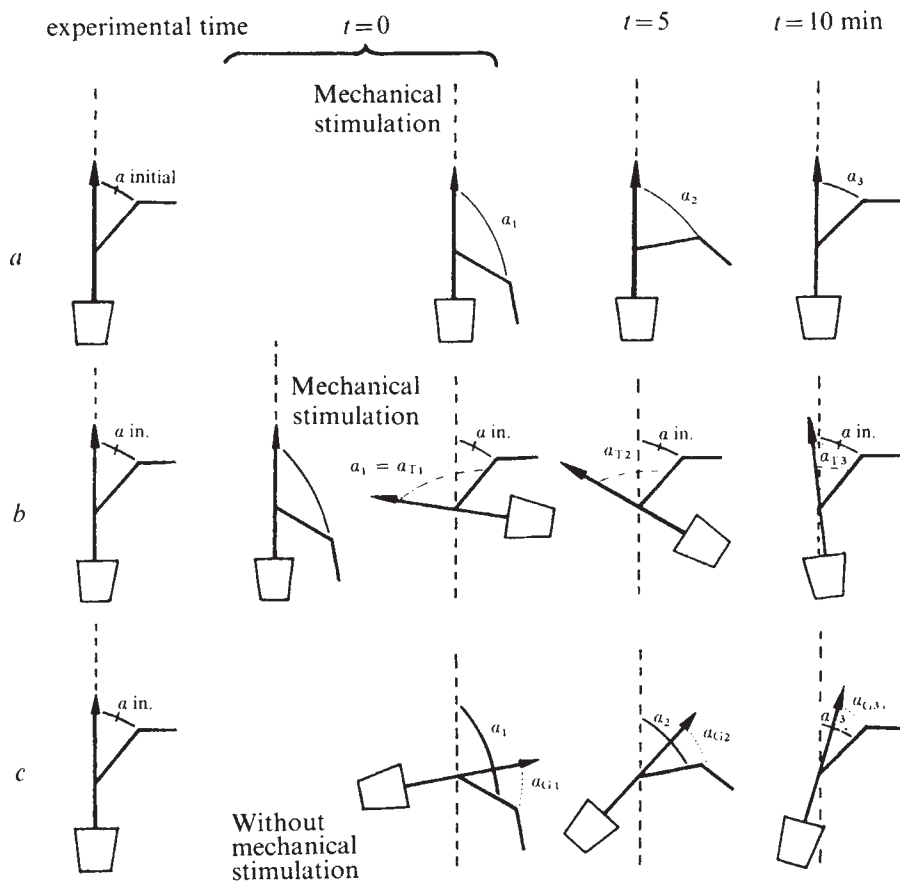


Fig. 2 Schematic representation of procedures used to determine the components of the recovery process of the primary pulvinus. *a*, Direct observation of recovery after mechanical stimulation quantified by angles α_1 and so on. This corresponds to Fig. 3*a*. *b*, Shock component measured by angles α_{T1} and so on when the geotropic effect is suppressed. This corresponds to Fig. 3*b*. *c*, Geotropic component alone, measured by angles α_{G1} and so on. This corresponds to Fig. 3*c*.

Diurnal periodicity can thus effectively be eliminated if all measurements are made during this part of the photoperiodic cycle, as was done during the investigation reported here. Phototropism was rendered negligible as diffuse fluorescent lighting was used. Geotropism, the tendency of part of a plant to orientate itself in a preferred direction with respect to gravity, is particularly marked in the sensitive plant¹⁰. If a leaf is displaced by carefully inverting the whole plant, it recovers its preferred orientation in 2 or 3 h. But when a leaf sinks due to mechanical stimulation, geotropism will still operate as the petiole has lost its preferred orientation. At the same time shock is a major factor.

To measure the recovery component due to shock alone, a leaf was stimu-

lated mechanically and geotropism was suppressed by maintaining the petiole at its initial orientation to the vertical throughout recovery by rotating the plant pot (Fig. 2*b*). The recovery component due to geotropism alone was determined as follows. First, I did a preliminary experiment with mechanical stimulation and with the plant pot fixed (Fig. 2*a*). Second, without mechanical stimulation, the pot was rotated each minute, to keep the angle between primary petiole and vertical the same as it was at the corresponding time in the preliminary experiment (Fig. 2*c*). These two procedures resulted respectively in curves (b) and (c) of Fig. 3. These two curves when added gave curve (d) which must be compared with the normal recovery (curve *a*). The agreement is adequate and, in

particular, the corresponding maxima and minima do not differ by more than 2 or 3 min during the first hour of observations. This analysis explains the three apparently different modes of recovery reported by Charnley *et al.*¹ in different leaves of this species.

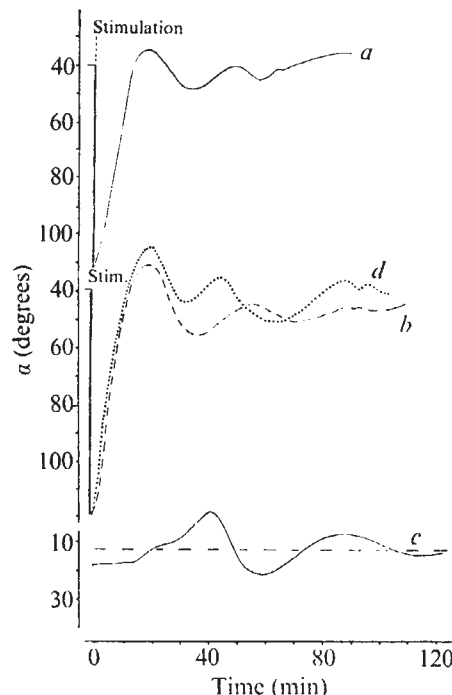


Fig. 3 Components intervening in the recovery process. *a*, Experimental temporal variation. *b*, Temporal variation due to seismonastic movement without geotropic effect. *c*, Temporal variation due to geotropism alone. *d*, Theoretical temporal variation obtained by adding curves (*b*) and (*c*).

These modes can be obtained by a simple relative shifting of curves (*b*) and (*c*), which has been proved experimentally.

I thank Professor P. Gavaudan, Drs Charnley and Perrin for advice and the CNRS for support.

GABRIEL ROBLIN

Laboratoire du Chronophytotron,
Station biologique de Beau-Site,
86000 Poitiers,
France

Crystal structure and Dirac's Large Numbers Hypothesis

Towe¹ has raised what he considers to be a serious difficulty against Dirac's² Large Numbers Hypothesis. He claims that geological evidence weighs heavily against multiplicative creation. We point out here two serious flaws in Towe's arguments which invalidate his conclusion.

Towe claims that the measured equality of lattice dimensions between a 3×10^9 -yr-old quartz crystal and a new quartz crystal weighs heavily against Dirac's multiplicative creation. Towe states that atomic distances in Dirac's theory vary like t^{-1} , t being cosmological time. Although it is true that atomic distances expressed in Einstein units do vary as t^{-1} , it is a conceptual error to use such a time dependence for comparison with measurement, since any observed value for distance is obtained with apparatus which measures only in atomic units.

In fact, Dirac's theory predicts that atomic distance scales formed from m (particle mass), e (particle charge), and h (Planck's constant) should not vary with t in atomic units, which are the ones to be used. Only the number of particles N , the gravitational constant G , and any quantity defined in terms of N and G (such as the radius of a star, or the orbit of a planet about a star) could depend on t when expressed in atomic units.

Even supposing that the lattice dimensions could vary with time, the only possible verification of this would be to perform the same measurement at two different times. The reason why measurements performed on two crystals at the same time must give identical results is simple. Both crystals are made of the same kinds of atoms with the same physical laws in effect—those laws which are valid at the time of measurement. Consequently, at any given time the lattice spacing for a given crystal structure must be the same regardless of whether the crystal is old or new; the past history of the crystal is irrelevant.

In point of fact, since lattice dimensions are expressed only in terms of m , e and h , which have no time dependence when expressed in atomic units (that is measurable units) Dirac's theory predicts that lattice dimensions do not change in time.

As mentioned repeatedly by Dirac^{3,4}, it is difficult to understand how the matter in very old rocks can have multiplied without disrupting their crystal structure. This is because N varies as t^2 ($M \propto t^2$, because $M = Nm$).

Towe's experimental result that both the molecular weight and X-ray density of unit cells of the two crystals are equal tends to support the opinion that the new matter is not interstitial. This in no way

constitutes a serious blow against Dirac's theory.

V. CANUTO
P. J. ADAMS
E. TSIANG

Institute for Space Studies, NASA,
New York, New York 10025

¹ Towe, K. M., *Nature*, 257, 115 (1975).

² Dirac, P. A. M., in *The Physicist's Conception of Nature* (edit. by Mehra, J.), 45 (Reidel, Dordrecht, 1973).

³ Dirac, P. A. M., *Pont. Acad. Commentarii*, 11, No. 46 (1973).

⁴ Dirac, P. A. M., in *Theories and Experiments in High Energy Physics* (edit. by Kursunoglu, B., *et al.*), 443 (Plenum, New York, 1974).

TOWE REPLIES—Canuto, Adams, and Tsiang¹ are absolutely correct in pointing out that my crystal lattice spacing argument² regarding Dirac's multiplicative creation model is invalid. The lattice spacings of old and new crystals will indeed always appear the same. Nevertheless I cannot agree that no setback to Dirac's multiplicative creation hypothesis (in its present form) is incurred by the observation that there is a reasonable agreement between X-ray and pycnometrically determined densities in the minerals of the oldest rocks. If new atoms cannot be added as interstitials then they must be added to new unit cells external to and part of the existing crystal structures. About 40% additional atoms are required for crystals 3×10^9 yr old. In accommodating the added unit cells into existing rocks, any extensive deformation with accompanying phase changes (recrystallisation) must be avoided to prevent resetting the radioactive 'clocks'. The alternative to interstitials would thus seem to transcend conventional physics because in addition to the multiplicative creation of atoms with time ($Nm \propto t^2$) it requires the simultaneous creation of the intercrystalline space ($V \propto t^2$) to accommodate them. If this is accepted as a realistic alternative, then a rock and all of its crystals would be getting larger, resulting in a secular trend toward increasing grain size in older rocks. What inconclusive data there are³ tend to support the opposite conclusion. Like Canuto's cosmological studies (in preparation), perhaps future petrographic studies can be designed specifically to deal with such an esoteric model of petrogenesis.

Smithsonian Institution
Washington, DC 20560

¹ Canuto, V., Adams, P. J., and Tsiang, E., *Nature*, 261, 438 (1976).

² Towe, K. M., *Nature*, 257, 115 (1975).

³ Nanz, R. H., Jr., *J. Geol.*, 61, 51 (1953).

¹ Charnley, T., Perrin, R., and Porter, D., *Nature*, 257, 389-390 (1975).

² Brunn, J., *Beitr. Biol. Pfl.*, 9, 307-358 (1909).

³ Ruge, U., in *Catalogue Films Scientifiques* (Service du film de recherche scientifique, Paris, 1947-1948).

⁴ Aimi, R., *Bot. Mag., Tokyo*, 76, 374-380 (1963).

⁵ Fondeville, J. C., thesis, Univ. Poitiers (1964).

⁶ Gavaudan, P., and Roblin, G., *C. r. Séanc. Soc. Biol.*, 165, 162-166 (1971).

⁷ Bose, J. C., in *The Motor Mechanism of Plants*, 429 (Longmans, Green, London and New York, 1928).

⁸ Wallace, R. H., *Am. J. Bot.*, 18, 288-307 (1931).

⁹ Roblin, G., thesis, Univ. Poitiers (1975).

¹⁰ Gavaudan, P., and Roblin, G., *C. r. Séanc. Soc. Biol.*, 163, 743-746 (1969).

reviews

Developmental biology

Peter Calow

Crystals, Fabrics and Fields: Metaphors of Organicism in Twentieth-Century Biology. By Donna Jean Haraway. Pp. 231. (Yale University: New Haven and London, March 1976.) £9.

As with the organismic system she describes, Miss Haraway's book is a tight-knit fabric of several interdependent parts. There are three main threads. The first is a lucid, well-researched account of the intellectual development of Ross G. Harrison, Joseph Needham and Paul Weiss. The second uses the work of these men to re-examine the time-honoured controversy between reductionism and holism. Finally these two themes are brought together in a critical appraisal of the Kuhnian idea of scientific progress. The strongest thread in Haraway's argument is the first, the weakest is the third, and for me the most interesting is the second.

It has become fashionable to dismiss the part-whole controversy as 'stone-cold dead', and yet, Miss Haraway effectively brings it alive again by showing how it has played a crucial rôle in the work of three of the most important and influential developmental biologists of this century. In this context she asks: "Is the organicism-reductionism controversy, in which Harrison and Co. were involved, the same as the vitalism-mechanism controversy of yester-year". Miss Haraway wishes to establish a clear distinction between these 'four isms' because she wishes to argue that they represent the alternative paradigms to which developmental biologists were subjected in the 1920s and 1930s. I will argue that, in several crucial respects, she does not succeed in this task and that in consequence she fails to establish, beyond doubt, the legitimacy of the Kuhnian interpretation.

With Miss Haraway I take mechanism to be, literally, the machine view of life and vitalism to be the involvement of indispensable, non-physical entities in biological explanation. Reductionism rests on the belief that the whole can be fully explained in terms of the parts whereas organicism (biological holism) asserts that the whole cannot be fully explained in this way. Mechanism is equivalent to reduc-

tionism if, and only if, machines are reducible. Alternatively it is easy to see how through Drieschian arguments from self-repair and Paley-Polanyi-type arguments from design, that mechanism may lead to holism and vitalism—not, however, by a series of discontinuous, paradigm changes, as implied by Miss Haraway, but simply by following the metaphor to its logical conclusion. Similarly, organicism, although not being equivalent to reductionism, may be compatible with it. One must know about the whole before one can reduce it in any meaningful sense of the word. Surely, Harrison's holism was operational throughout; Needham began in this way but then progressed to a more uncompromising view of wholeness; Weiss seems always to have had anti-reductionist leanings. The discontinuities, if they exist at all, are most clear in the development of Needham. In spite of Haraway's lucid factual account, however, paradigm changes are not at all self-evident in the work of any of these men.

Haraway, as do the organismic biologists themselves, rejects the idea that organicism is equivalent to vitalism. And yet if one accepts an evolutionary cosmogeny in which whole is formed from parts, it is difficult to understand how properties could have emerged in the wholes which were not explicable in terms of the constituent parts. Otherwise as Vernon Pratt has argued in an important paper not cited by Miss

Haraway (*Stud. Hist. Phil. Sci.*, 5, no. 1, 1-13, 1974), it is necessary to postulate the coming into being of extra inexplicabilities between levels. But from whence do these inexplicables arise—like Aristotle's *Pneuma*, from an extraterrestrial Quintessence? I suggest that there is in fact little difference between organicism in this extreme, non-operational form and the old vitalism, and that the former has evolved out of the latter.

Miss Haraway suggests that the strength of the Kuhnian approach is its concern for the extralogical, emotional aspects of science. This is also its weakness, for it is difficult, if not impossible to judge the paradigm idea in any dispassionate objective sense, and it is all too easy to make the facts fit the model. If revolution, discontinuity, crises, and so on, have been such important elements in the scientific advance, they should surely emerge as self-evident processes from the history of science. They do not emerge in this way from *Crystals, Fabrics and Fields*. In spite of this criticism, however, the book is still worth reading for its rich factual content, and were it not for the price it could provide a valuable adjunct to undergraduate, developmental biology courses. □

Dr Calow is a lecturer in the department of zoology at the University of Glasgow, UK.

Sedimentation within the tidal zone

Tidal Deposits: a Casebook of Recent Examples and Fossil Counterparts. Edited by Robert N. Ginsburg. Pp. xiii+428. (Springer: Berlin and New York, 1975.) DM 85; \$34.80.

THIS volume, dedicated to the memory of Rudolf Richter for his pioneer research on tidal deposits, is an outstanding compilation of current knowledge of sedimentation within the tidal zone. Forty-five papers on selected recent and fossil tidal deposits are contributed by 52 authors, and it is refreshing to discover clear evidence of editorial authority in achieving a uniform style of presentation of the papers. Robert

Ginsburg, in liaison with George de Vries Klein, John Harms and Richard Boersma, supplied authors with a definitive, itemised format, even adopting standard symbols for the preparation of facies diagrams. Thankfully, authors responded well in producing what are essentially a series of factual summaries of the principal facies types, vertical facies sequences, sedimentary structures, and faunas and floras characterising the various tidal deposits. All papers are copiously illustrated with photographs and invaluable graphical logs and facies maps. Consequently, this casebook provides sedimentologists with the necessary criteria for recognising and readily comparing sedimentary se-

quences deposited within tidal environments.

Two major subdivisions are considered—the siliciclastic and carbonate tidal deposits. Recent siliciclastic examples include the classic tidal flats of the North Sea, Mont Saint-Michel Bay and the Bay of Fundy, transgressive tidal deposits in barrier island complexes of the Atlantic coast of North America, fine-grained deposits on the margins of the Colorado River Delta, wind-tidal flats of Laguna Madre, and deposition within tidal inlets at Long Island and the Dutch coast. Twelve fossil analogues are described, ranging in age from Precambrian to Pliocene, but mostly from the Palaeozoic of North America. The examples vary from those showing many of the diagnostic sedimentary features of a

North Sea model, to a type in which the sedimentary record shows little evidence of sea-level fluctuation.

Recent carbonate examples are naturally from the Bahamas, Arabian Gulf and Shark Bay, Western Australia, for which subtidal, intertidal and supratidal deposits are described. The 19 fossil counterparts are discussed in three sections—examples showing the three tidal zones of accumulation; those characterised by laminated and stromatolitic facies; and a miscellany illustrating how it is possible to recognise ancient tidal deposits on the basis of sedimentary features and facies pattern.

This unique casebook ends with an annotated bibliography of what are regarded as the 25 key papers on tidal sedimentation. **Brian Waugh**

Form from top to bottom

Orientation and Form. By Irvin Rock. Pp. 165. (Academic: New York.) £5.75.

THIS book is the outcome of twenty years research: it addresses two questions. First, what determines the orientation in which we see a visual form—that is, which part of it do we take to be the top? It could be the position of the form on the retina, its orientation with respect to gravity, or its orientation with respect to the visual framework. Usually all three factors coincide, but Rock has explored systematically what happens when we view forms with head tilted or inverted and when there is no visual framework. He has also shown that there are three further determinants of what we take to be the top of a shape—familiarity with the shape in a given orientation, the endogenous properties (such as asymmetry) of the shape, and instructions. Rock gives a most convincing account of how these factors interrelate and of what determines which will predominate when they are in conflict.

The second question Rock considers is what determines whether a shape will be recognised when its orientation is changed, and here he successfully resolves two apparent contradictions in the literature. It is normally found that when the head is tilted, forms are seen the way up they are in space not the way up they are on the retina. Nevertheless, certain kinds of complex but familiar forms, like faces and letters, become very hard to recognise when inverted on the retina. He argues that it takes time to correct for

changes in orientation on the retina caused by head tilt and that the correction process operates in a piece-meal fashion: it is therefore difficult to correct fully when a complex shape such as a face is presented in an unusual orientation. He cites experimental evidence of his own, much of which has not previously been published, to support his claim. He also gives a convincing explanation of why it is that some forms are most difficult to recognise when they are turned through 180°, others when they are turned through intermediate angles. He anticipates some of Shepard's more recent (and currently better known) work on the problem of orientation: in many ways Rock's insights go deeper and take into account a wider range of factors.

Orientation and Form is modest both in length and manner. It is an excellent example of how an approach based on common sense and carefully designed experiments can resolve important issues and lead to the construction of a convincing theory that is anything but obvious to common sense—except in retrospect.

N. S. Sutherland

Success in ion transport

Ion Transport in Plant Cells and Tissues. Edited by D. A. Baker and J. L. Hall. Pp. xiv+437. (North-Holland: Amsterdam and Oxford; American Elsevier: New York, 1975.) Dfl. 130; \$54.25.

THE editors of this book have dealt successfully with a difficult problem confronting those who assemble a

multi-author volume; the level and tone of address to the reader is relatively uniform. To have achieved this while preserving the individual styles of writing indicates how much unobtrusive work must have been done.

The level of address is fairly advanced, making the book most valuable to postgraduate students, research workers and teachers seeking an entry into the detailed literature. In several chapters there are summaries of the state of knowledge in the form of schemes and models, and there are useful compilations of information in tabular and chart form.

There is insufficient emphasis in many of the chapters on how experiments are carried out and not enough comment on the reliability of techniques. The experimentally-minded reader would have been helped more if the frailties of certain procedures had always been dealt with as straightforwardly as in excellent chapters by R. G. Wyn Jones (Ion Uptake in Excised Roots) and J. L. Hall (Cell Membranes).

In their introductory chapter Baker and Hall reproduce a most thought-provoking illustration, from a book compiled in 1959 by F. C. Steward, in which salt accumulation is related to overall processes of growth and metabolism. Nearly 20 years on it is disappointing how little attention several of the authors are prepared to give to this integrated approach. It is surely this reluctance, coupled with experimental convenience, that packs the book so full of information about the absorption of chloride, simply because it is *not* involved to any major extent with metabolism. It is regrettable that the coverage given to the published work on the absorption of the nitrogenous ions, and on phosphate and sulphate in relation to their metabolism is far from complete.

The book ends with three enjoyable chapters in which ion transport is related to specific functions (T. J. Flowers on Halophytes; U. Lüttge on Salt Glands; D. A. Thomas on Stomata). These, taken with two splendid chapters by J. B. Hanson and D. E. Koeppe on Mitochondria and P. S. Noble on Chloroplasts, suggest that it might have been better to have adopted this functional approach throughout the book rather than, in the middle chapters, to have described ion transport in higher plants according to the experimental approach by which it is studied.

The cost of this book, for all its merits, is a staggering \$54.25. It must be a great disappointment to the editors that at such a price, the book will be beyond the financial reach of most individuals and many smaller libraries.

D. T. Clarkson

obituary

Raymond Bridgeman Cowles, Professor of Zoology at the University of California, Los Angeles, died on December 7, 1975, at the age of 79. Born in Natal, South Africa, he spent his childhood there but was educated in the United States, receiving his doctorate at Cornell University in 1928.

Two themes from his Natal boyhood were stamped on Cowles' subsequent career: first, the many animal-environmental relationships he saw in the bush near his home in the Umzumbe Valley, such as the monitor lizards that laid their eggs in the almost impregnable fortresses of termite nests, and second, the transformation of the Zulu to a state of deep poverty as their numbers increased. His career was further stimulated by the work he did to raise money for his education: he worked as a 'muleskinner' and as a canal guard in the then desert wilderness of the Imperial Valley of California.

His tools were the simplest of instruments: the camera, the thermocouple and thermometer, and his own critical powers of observation, yet he founded the field of thermal relationships of reptiles. He did more classic work on the pit-organ of rattlesnakes, where he demonstrated the exquisite thermal sensitivity of this receptor. Blindfolding the rattlesnake, he swung before it a glowing electric bulb encased in black cloth. Often enough, the snake struck at the "warm prey object", hitting it in midswing with a deadly-accurate spray of venom. One had no trouble envisioning a leaping desert mouse or pocket gopher receiving the same treatment, all through central

integration of afferent impulses from the most sensitive thermal receptor in the animal kingdom.

The same quality of insight led to the remarkable generalisation that male germ cells in diverse species are inactivated by exposure to approximately the same thermal levels. Central to this finding were observations that male reproductive tissues are protected from high temperatures in various ways; for example, the scrotum of mammals and the placement of gonads of birds adjacent to air sacs within the body. Cowles also saw that blood flow was responsive to temperature, and that this dynamic interaction was crucial to thermoregulation in reptiles, birds, and mammals, proving this by beautifully 'simple' experiments.

Professor Cowles also performed early work on water conservation and loss in desert animals and many other ecological subjects. Another dominant factor in his later career was his realisation that the human race violated the rules by which living things in nature subsisted in ecological harmony. A major book on this subject was just completed at the time of his death and awaits publication.

An important part of Cowles' writing was for the interested public, but he taught most of all by encouraging his students to participate in the excitement and insights of field trips, thinking with him as he pondered on what he saw. A motor trip through the desert might find him ranging across topics as diverse as the migration of waterfowl over established desert flyways, the economic use of the jojoba plant (which is now being used

to provide a substitute for sperm whale oil), the dynamics of soil and sand movement in response to wind and water, the basking behaviour of rattlesnakes on warm pavements at dusk, or the aerodynamics of bird flight—the car's speed would be adjusted to that of a bird flying alongside, and its speed checked.

Professor Cowles was serenely unimpressed by his many academic accolades, and everywhere he was much respected for his deep honesty and warmth as a human being.

Kenneth S. Norris
Clara M. Szego

John Buchan Jepson, who died on May 7 aged 53, was the Professor of Biochemistry at the Middlesex Hospital Medical School.

After taking a first degree at Cambridge, he went to work for his doctorate at Oxford, where he did research on Penicillin under the supervision of Sir Robert Robinson. He later worked on the biochemistry of amino-acids, hydroxylases and psychomimetic amines, and also on hereditary diseases and chromatography. He was also very interested in Biochemical Education; he was always trying to find new methods of conveying the excitement, importance and relevance of his subject—and as enthusiasts always do, he succeeded in involving his audience. Another aspect of this interest was the efficient and rapid communication of research. He was chairman of the Committee on Biological Information, and worked right up till his death on new means of data dissemination.

announcements

Meetings

June 10, **Conservation of Wild Life in Kenya**, London (The Secretary, The Royal Society of Arts, John Adam Street, London WC2 6EZ).

July 5–8, **Mathematics of Hydrology and Water Resources**, University of Lancaster (The Secretary and Registrar, Institute of Mathematics and its Applications, Maitland House, Warrior Square, Southend on Sea, Essex SS1 2JY, UK).

September 6–9, **Nonlinear Waves and**

Solitons, University of Newcastle upon Tyne (The Secretary and Registrar, Institute of Mathematics and its Applications, Maitland House, Warrior Square, Southend on Sea, Essex SS1 2JY, UK).

September 14–16, **Theoretical Developments in Control**, University of Leicester (The Secretary and Registrar, Institute of Mathematics and its Applications, Maitland House, Warrior Square, Southend on Sea, Essex SS1 2JY, UK).

December 21–22, **The Structure and**

Assembly of Biological Systems, London (Dr A. C. Allison, Clinical Research Centre, Watford Road, Harrow, Middlesex, UK).

April 18–22, 1977, **Fifth International Conference on Reactor Shielding**, Knoxville, Tennessee (Deadline for abstracts: August 20) (D. K. Trubey, Radiation Shielding Information Centre, Oak Ridge, National Laboratory, PO Box Y, Oak Ridge, Tennessee 37380).

July 11–15, 1977, **Veterinary Parasitology**, Sydney, Australia (The Secre-

tary, 8th International Conference on WAAVP, CSIRO, McMaster Laboratory, PMB No 1, Post Office, Glebe N.S.W. 2037, Australia).

August 15-18, 1977, **Chemical Thermodynamics**, Ronneby, Sweden (Fifth International Conference on Chemical Thermodynamics, Chemical Centre, Lund University, PO Box 740, S-220 07 Lund, Sweden).

September 26-28, 1977, **On-Line Surveillance and Monitoring of Process Plant**, London (Deadline for abstracts: October 31) (Conference Secretariat, On-Line Surveillance and Monitoring of Process Plant, Society of Chemical Industry, 14 Belgrave Square, London SW1X 8PS, UK).

Reports and publications

Other countries

World Health Organization. A Multi-Axial Classification of Child Psychiatric Disorders: an Evaluation of a Proposal. By Michael Rutter, David Shaffer and Michael Shephers. Pp. 85. Sw.fr.18. WHO Food Additives Series, No. 7: Specifications for the Identity and Purity of Some Food Colours, Flavour Enhancers, Thickening Agents, and Certain Food Additives. Pp. v + 216. Alternative Approaches to Meeting Basic Health Needs in Developing Countries. Edited by V. Djukanovic and E. P. Mach. (A Joint UNICEF/WHO Study.) Pp. 116. Sw.fr.24. (Geneva: WHO; London: HMSO, 1975 and 1976.) [242]

World Health Organization. International Agency for Research on Cancer: Annual Report, 1975. Pp. 149. (Lyon: International Agency for Research on Cancer; London: HMSO, 1975.) Sw.fr.12. [252]

World Health Organization. Public Health Papers, No. 64: Health Economics—Report on a WHO International Seminar. Pp. 44. (Geneva: WHO; London: HMSO, 1975.) Sw.fr.6. [262]

United States Department of the Interior: Geological Survey, Professional Paper 823-A, B, C, D: Paleozoic Corals of Alaska. By Michael Churkin, Jr., et al. Pp. 75 + 28 plates. (Washington, DC: US Government Printing Office, 1975.) [262]

The Dr. H. P. Heineken Prize. By C. Bauer. Pp. 32. (Amsterdam: The Heineken Foundation, 1975.) [272]

Environment Canada. Fisheries and Marine Service. Technical Report No. 520: Two Computer Programs for Analysis of Catch-at-Age Data. By W. G. Doubleday. Pp. 7 + POPI Listing. (St. Andrews, NB: Research and Development Directorate, Biological Station, 1975.) [272]

Food and Agriculture Organization of the United Nations. Cereal Seed Technology: a Manual of Cereal Seed Production, Quality Control and Distribution. Edited by Walther P. Feistritz. (FAO Agricultural Development Paper No. 98.) £2.60. Population, Food Supply and Agricultural Development. Pp. v + 62. 80p. Report of the Fifth Session of the FAO Panel of Experts on Integrated Pest Control, held in Rome, Italy, 15-25 October 1974. Pp. 41. 80p. Land Reform: Land Settlement and Co-operatives. No. 1/2, 1974. Pp. 147. £1.60. (Rome: FAO; London: HMSO, 1975.) [23]

CERN—European Organization for Nuclear Research. CERN 75-19: Instruction Timing for the CDC 7600 Computer. By H. Lipps. Pp. 73. (Geneva: CERN, 1975.) [23]

Annals of the South African Museum, Vol. 68: Monograph on the Hydroidea of Southern Africa. By N. A. H. Millard. Pp. 513. (Cape Town: South African Museum, 1975.) R.22.50 [43]

Cultivo de la Fresa en el Laru. (Boletín Técnico, No. 80). Pp. 35. (Lima: Ministerio de Agricultura, 1974.) [43]

Smithsonian Contributions to Zoology, No. 214: Myodocopid Ostracoda from Southern Africa. By Louis S. Kornicker. Pp. 39. (Washington, DC: Smithsonian Institution Press, 1976. For sale by US Government Printing Office.) [43]

US Department of the Interior: Geological Survey, Professional Paper 779: The Story of the Water Supply for The Comstock, Including the Towns of Virginia City, Gold Hill, and Silver City, Nevada, Together with Other Water-Related Events for the Period 1859-1969. By Hugh A. Shamberger. Pp. xi + 53. (Washington, DC: US Government Printing Office, 1972.) 70 cents. [43]

US Department of Agriculture. Farmers' Bulletin No. 2200: Controlling Tomato Diseases. Pp. 12. (Washington, DC: Government Printing Office, 1975.) 30 cents. [53]

US Department of Agriculture. Agricultural Research Service. Technical Bulletin No. 1521: Host Records for Lepidoptera Reared in Eastern North America. By D. C. Ferguson. Pp. 49. (Washington, DC: Government Printing Office, 1975.) [53]

Australia: Commonwealth Scientific and Industrial Research Organization. CSIRO Twenty-seventh Annual

Person to Person

A self-contained one-bedroom apartment in a modern building, conveniently located for central and West-End London by public transport; central heating, carpeted throughout. Exchange for apartment or house in Toronto-Hamilton area. Dr L. Knight, 370 Dixon Road, Apartment 1004, Weston, Ontario M9R1T2 Canada.

Science and Music. Physics librarian compiling a comprehensive bibliography on the collaboration between science and music, from the 16th to 20th centuries, would like to exchange information with other persons interested in the subject. Information relating to unorthodox theories and contraptions is specially desired. Please write to James R. Parrott, c/o EMS Library, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1.

All relevant material on the biological and ecological effects of extremely low frequency radiation are being sought by the NRC committee studying the US Navy's Project Seafarer.

All material and references should be submitted in writing as soon as possible to Samuel Abramson, Committee on Biosphere Effects of Extremely Low Frequency Radiation, National Research Council, 2101 Constitution Avenue, N.W., Washington, D.C. 20418.

The Fleet Study Group, is interested in the formation and structure of the Chesil Beach; the sediment of the floor of the Fleet; the physico-chemical characteristics of the Fleet water and the biological communities of the Fleet. Anyone who can provide information on these topics or who wishes to help the study, please contact the Secretary, Mrs J. M. Fitzpatrick, c/o College of Education, Cranford Avenue, Weymouth, Dorset DT4 7LQ.

There will be no charge for this service. Send items (not more than 60 words) to Martin Goldman at the London office. The section will include exchanges of accommodation, personal announcements and scientific queries. We reserve the right to decline material submitted. No commercial transactions.

Report, 1974/1975. Pp. 75. (Melbourne: CSIRO, 1976.) [83]

World Meteorological Organization. Technical Notes. No. 142: Applications of Satellite Data to Aeronautical Meteorology. By F. G. Finger and R. M. McInturff. Pp. 93. No. 144: Rice and Weather. By G. W. Robert-

son. Pp. 40. No. 145: Economic Benefits of Climatological Services. By Dr. R. Berggren. Pp. 43. (Geneva: WMO, 1975.) [83]

New Health Practitioners. (A Conference Sponsored by The John E. Fogarty International Center for Advanced Study in the Health Sciences and the Association of the Teachers of Preventive Medicine, National Institutes of Health, Bethesda, Md., May 14-15, 1974.) Pp. xvi + 156. (Bethesda, Md.: US Department of Health, Education and Welfare, Public Health Service, National Institutes of Health, 1975.) [83]

Bulletin of the Auckland Institute and Museum, No. 10: Taxonomic Revisions in the Family Haloragaceae. 1. The Genera Haloragis, Haloragodendron, Glischrocaryon, Meziella and Gonocarpus. By A. E. Orchard. Pp. 299. (Auckland: Auckland Institute and Museum, 1975.) [83]

Geological Survey of Western Australia. Annual Report for 1974. Pp. 112. Memoir No. 2: The Geology of Western Australia. Pp. 541. (Perth, W. Australia: Geological Survey of W.A., 1975.) [83]

National Research Council Canada. Annual Report on Scholarships and Grants in Aid of Research, 1974-75. Pp. xiii + 574. (Ottawa: National Research Council Canada, 1976.) \$2.50. [93]

Det Norske Videnskaps-Akademi. I. Mat.—Naturv. Klasse. Skrifter. Ny Serie. No. 34: Geochemistry of the Oslo Basaltic Rocks. (Studies on the Igneous Rock Complex of the Oslo Region, XXIV.) By Peter W. Weigand. Pp. 38. No. 35: Radioelement Distribution and Heat Production in Precambrian Granitic Rocks, Southern Norway. By P. G. Killen and K. S. Heier. Pp. 32. (Oslo, Bergen and Tromsø: Universitetsforlaget, 1975.) [103]

Tobacco Research, Vol. 1, No. 1, June, 1975. Pp. 1-88. Annual subscription Rs. 30 in India; \$10 or £4 overseas. \$2 or £1 extra for Airmail. (Rajahmundry, India: Tobacco Research Workers' Association of India, 1975.) [103]

Gt Britain

Heriot-Watt University. Annual Report, 1974-75. Pp. 84. (Edinburgh: Heriot-Watt University, 1976.) [13]

The Carnegie United Kingdom Trust. Sixty-second Annual Report, 1975. Pp. 30. (Dunfermline, Fife: The Carnegie United Kingdom Trust, 1976.) [13]

Public Service Broadcasting: A European View. By Albert Scharf. (BBC Lunch-Time Lectures, Tenth Series, 3.) Pp. 13. (London: BBC, 1975.) [13]

International Journal of Hydrogen Energy, Vol. 1, No. 1, 1976. Pp. 1-108. Subscription rates for 1976: Libraries, Research Establishments, and all other multiple-reader institutions \$75. Personal subscription \$55. (Oxford and New York: Pergamon Press, 1976.) [13]

Memoirs of the Royal Astronomical Society, Vol. 81, Part 1: Numerical Simulation of Natural Photometric Systems. By A. W. J. Cousins and D. H. P. Jones. VRI Standards in the E Regions. By A. W. J. Cousins. Pp. 36. (Oxford and London: Blackwell Scientific Publications, 1976. Published for The Royal Astronomical Society.) [13]

National Radiological Protection Board. NRPB R 37: Thermoluminescence Sensitivity Variations in LiF: PTFE Dosimeters Incurred by Improper Handling Procedures. By E. W. Mason, A. F. McKinlay, I. Clark and D. Saunders. Pp. 10. (Harwell, Didcot, Oxon: National Radiological Protection Board, 1975.) [13]

Biochemical Engineering: An Industrial Fellowship Report for The Science Research Council and The Institution of Chemical Engineers. By A. N. Emery. Pp. 47 + vii + viii. (London: Science Research Council, State House, High Holborn, 1976.) gratis. [13]

Imperial Metal Industries, Ltd. Annual Report for 1975. Pp. 40. (Birmingham: Imperial Metal Industries, Ltd., 1976.) [13]

University of Cambridge. Local Examinations Syndicate. Examinations at 16+—Comments on the Report to the Schools Council. Pp. 34. (Cambridge: The University, Local Examinations Syndicate, 1976.) [23]

Social Science Research Council. Progress and Problems in Social Forecasting. Edited by Christopher Freeman, Marie Jahoda and Ian Miles. Pp. 85. (London: Social Science Research Council, 1976. Available from Research Publications Services, Victoria Hall, Fingal Street, London, SE10.) £1.25 + 14p handling charge. [23]

Department of the Environment. Water Pollution Research 1973: Report of the Director of Water Pollution Research. Pp. vi + 144 + 4 plates. (London: HMSO, 1976.) £2.55 net. [33]

Metrickation Board. Going Metric—Farming and Horticulture. Pp. 21. (London: Metrickation Board, 1975.) [43]

Philosophical Transactions of The Royal Society of London. B: Biological Sciences. Vol. 274, No. 928: Recent and Prospective Trends in Fertility in Developed Countries. By D. V. Glass. Pp. 1-52. (London: The Royal Society, 1976.) UK £2.85; Overseas £2.95. [43]

Radioactive Low Energy Photon Sources. (Technical Bulletin 75/6.) Pp. 20. (Amersham: The Radiochemical Centre, 1975.) [53]

Bulletin of the British Museum (Natural History). Botany, Vol. 5, No. 4: A Revision of the Macaronesian Genus *Argyranthemum* Webb Ex Schultz Bip. (Compositae-Anthemideae.) By C. J. Humphries. Pp. 145-240 + plates 28 and 29. (London: British Museum (Natural History), 1976.) £6.50. [83]

The Craftsman in Engineering: An Interim Report. By Muriel Venning, Owen Frith and Chris Grimbley. Pp. 39. (London: Engineering Industry Training Board, 1975. Available from EITB Publications Department, P.O. Box 176, 54 Clarendon Road, Watford.) £1.05 (remittance with order.) [93]